LECTURES ON THE
SCIENTIFIC BASIS OF MEDICINE
1956-57

British Postgraduate Medical Federation
University of London

LECTURES ON THE SCIENTIFIC BASIS OF MEDICINE

Volume VI
1956–57

UNIVERSITY OF LONDON
THE ATHLONE PRESS
1958

Published by
THE ATHLONE PRESS
UNIVERSITY OF LONDON
at 2 Gower Street, London WC1
Distributed by Constable & Co Ltd
12 Orange Street, London WC2
Canada
University of Toronto Press
U S A
Essential Books Inc
Fair Lawn N J

© *British Postgraduate Medical Federation* 1958

Printed in Great Britain by
WESTERN PRINTING SERVICES LTD
BRISTOL

PREFACE

THE sixth annual volume of *Lectures on the Scientific Basis of Medicine* includes twenty one of the twenty nine lectures delivered during the winter of 1956–57 These lectures are arranged annually by the British Postgraduate Medical Federation and are designed for graduates in medicine looking forward to careers as clinical teachers and consultants in medicine and surgery and their special branches or to careers in research in the medical sciences

The lectures in this volume deal for the most part with subjects that carry further the problems dealt with in the earlier volumes, and reflect the use of new techniques and the trends of modern research Thus there is again a number dealing with biochemical and metabolic problems and the action of enzymes and others that apply the methods of the physicist to biological problems The former are represented by the lectures by Long on Phospholipids, Greenbaum on The Control of Fat Metabolism, Whelan on The Synthesis and Degradation of Polysaccharides, and by Sinclair on Essential Fatty Acids in Nutrition' and the latter by the lectures by Melrose on the Principles of Heart Lung Machines Gilson on Lung Function Tests', and Smith on The Electron Microscope in the Study of Viruses The centenary year of the birth of Sherrington is reflected in Liddell s lecture on Cajal and Sherrington and also in a group of lectures dealing with application of physico chemical methods to the study of the nervous system, viz McIlwain on Neurochemistry Whittaker on Metabolism of Acetylcholine in Nervous Tissues, Zaimis on Neuromuscular Blocking Substances, and Keele on Causes of Pain The metabolic action and fate of hormones are gaining increasing recognition as important factors in health and disease and these are dealt with in two lectures one by Hadfield on the Endocrine

System in Breast Cancer', the other by Myant on 'Biliary Secretion of Thyroid Hormones' The preservation of healthy living tissues outside the body and their substitution for diseased structures in patients is a subject of practical importance to surgeons as well as biologists, and three lectures illustrate the advances in this field, by Fell on 'The Physiology of Skeletal Tissue in Culture,' Dempster on 'Homo-transplantation of Organs', and Robb on 'Arterial Substitutes'

In addition to these lectures that can be regarded as continuing and developing subjects and methods dealt with in earlier years, this volume includes four on problems of immediate practical interest and new methods of investigation Sir John Cockcroft's lecture on 'The Biological Significance of Atomic Energy' suggests many new fields that may require medical investigators trained in new ways, as does also Lewis' lecture on 'Social Psychiatry', while Barber's on 'Resistance of Staphylococci to Antibiotics' and Seddon's on 'Poliomyelitis' illustrate new methods of attack on urgent problems

The lectures delivered during the winter of 1957–58 are now being prepared for publication in a further volume, Volume VII of the series, and will include many that, in addition to illustrating continued progress in subjects and by methods already discussed in the earlier volumes, will stress how the study of the cell, its structures and their functions is becoming possible and necessary for a better understanding of health and disease

FRANCIS R FRASER
Director British Postgraduate Medical Federation

5 February 1958

CONTENTS

NOTE

The lectures printed in this volume were delivered on the following dates

I	11 October 1956	XI	23 October 1956
II	6 November 1956	XII	24 January 1957
III	21 February 1957	XIII	29 November 1956
IV	8 November 1956	XIV	14 February 1957
V	22 November 1956	XV	7 February 1957
VI	27 November 1956	XVI	19 February 1957
VII	17 January 1957	XVII	10 January 1957
VIII	15 January 1957	XVIII	5 February 1957
IX	29 January 1957	XIX	25 October 1956
X	22 January 1957	XX	12 February 1957

XXI 16 October 1956

I

The Biological Significance of Atomic Energy

SIR JOHN COCKCROFT

THE post war developments of atomic energy have very important biological implications The facilities provided by nuclear reactors for producing radioactive isotopes in virtually unlimited amounts have provided a major new research tool for the biologist, particularly the biochemist, whilst the new sources of radioactivity have greatly helped the radio-therapist On the other side of the picture there are the potential hazards which might arise in the widespread use of radioactive material and the development of nuclear power I will try to show that these hazards can be controlled by the scientific approach to radiation protection which has been adopted

Nuclear reactors today produce about twenty two radioactive isotopes for biological research Amongst these the most important are Tritium, Carbon 14, Sodium 22 and 24 Phosphorus 32, Sulphur 35 Iron 55 and 59 and Cobalt 60

Carbon 14 is produced by irradiating nitrogen in a reactor Aluminium nitride is placed in the Windscale reactors for a period of about a year or more during which time about 300 mc/year of carbon 14 are produced from each kilogram of aluminium nitride The carbon 14 is then extracted at the Radio chemical Centre Amersham, and is usually incorporated into one of a number of componds required by the biologist In some cases the classical methods of the organic chemist are used In other cases biological synthesis is used For example, tobacco

plants are used to synthesize glucose and starch, a diligent hen was fed on chlorella labelled with carbon 14, the hen then produced egg albumen labelled with carbon 14 About 180 compounds are now available

The Uses of Atomic Energy in Biological Research

These radioactive isotopes have been applied to a very wide range of problems in biological research where they have provided a new tool of comparable importance to chromatography As a physicist, I am not competent to describe their applications in detail, I can only give some general impressions gathered mainly at the Geneva conference on the peaceful uses of atomic energy

The radioisotopes are of course used as tracers The great sensitivity of the modern methods of detection of radioactivity allows their passage through the organisms to be traced as they pass from one molecular form to another just as though the tracer atom carbon 14, or iron 59 was visible to the experimental biologist throughout its passage from intake to excretion

A general result of these studies has been to change in a major way earlier ideas on metabolism On the old ideas, food was thought of as a fuel to be used for supplying energy or repairing wastage in tissues A major attack on this viewpoint came from the work of Schoenheimer using stable isotope tracers In summarizing his results he has said that the large complex molecules in food and the less complex compounds which form them —the fatty acids amino acids and nuclear acids—continually participate in rapid chemical reactions Ester peptide and other kinds of bonds are broken The fragments of complex molecules so released are mixed with those from other similar molecules and with those absorbed from the alimentary canal, forming a metabolic reserve Some of the fatty acid molecules are completely destroyed others of similar structure are continually formed from all kinds of substances and in particular from carbohydrates Similar reactions also take place with the decaying products of proteins Thus the biological system is a vast cycle of closely interrelated chemical reactions'

As a result of this work it has been realized that many simple substances, such as acetate, carbon dioxide, glycine and related compounds have great importance in biological economy

The research work has also demonstrated the rapid rate of turnover of particular molecules Thus amino acids are assembled into proteins in about 25 minutes Studies with radioactive iron e g Hevesy, show that of the 4 milligrams of iron present in human blood plasma, half of it leaves the circulation in about 90 minutes Half of this finds its way into the bone marrow to be used for the formation of haemoglobin of the blood corpuscles So 18 milligrams a day is used for haemoglobin formation The other half finds its way to the liver and other organs In cases of refractory anaemia the iron atoms disappear less rapidly from the plasma, so studies on anaemia are possible

Studies using radioactive sodium show that half the sodium in the blood plasma leaves the plasma in one to two minutes Studies with radio-carbon have shown that even the crystalline structure of the bones is in a state of flux It has also been possible to work out the exact metabolic pathways by which important compounds have been made Thus the origin of every carbon atom in the complex porphyrin molecule has been determined, the pathways of formation of the blood pigment haematin have similarly been unravelled

A recent paper from Russia by Grodzensky described the applications of isotopes to research on the pathogenesis of metabolic diseases Thus the metabolism of uric acid in healthy and gout afflicted persons was studied Healthy and ill persons were injected intravenously with uric acid marked with ^{15}N The metabolic reserve of uric acid which in a healthy person is about one gram, is increased by a factor of 15 in gout The origin of the increased production of uric acid was found by oral administration of glycine marked with ^{15}N to healthy and sick persons In the sick persons the uric acid was synthesized at three or four times the rate found in healthy persons The establishment of these facts makes it possible for the investigator to seek rational methods of therapy for this disease—directed towards reducing the synthesis of uric acid in the organism

The general opinion at Geneva was that this new knowledge

and change of concept was leading to important *revisions in the* search for therapy of metabolic diseases However, only a beginning has been made—mainly by describing the metabolic processes—we are apparently still far from understanding in detail how they work It remains for the future to reveal how the new knowledge can be used for the conquest of disease'

In plant physiology the use of carbon 14 as a tracer has been of great value in the study of the mechanism of photosynthesis—the process by which plants take in carbon dioxide together with water to produce carbohydrates and oxygen The experiment consists of adding carbon dioxide containing carbon 14 to unlabelled carbon dioxide and this is absorbed by algae After quite a short time—10 seconds or 60 seconds for example—the plant is killed by sudden treatment with boiling ethanol, so that all enzymatic processes are halted Extracts of the plant material are made and two dimensional chromatography is used to study the synthesis of molecules A drop of the solution is put on to the corner of a large filter paper The solution then diffuses down and the different compounds are separated The filter paper is then turned through a right angle and the compound again diffuses down so that a two-dimensional separation is achieved A photographic film is then placed against the filter paper and any molecules which contain carbon 14 blacken the film as in Plate I Figures 1 and 2 [1] After 10 seconds phosphoglyceric acid is predominant so that it must be the first compound into which carbon dioxide is built By the end of 60 seconds sugar phosphates and acid phosphates have appeared The process of build up of the carbohydrates, proteins and fats is an extremely complex one and a great deal of light has been thrown on the successive steps in the reaction

Radioactive isotopes are also used as sources of radiation for the treatment of malignant disease Twenty years ago radium was used in quantities of several grams for radio-therapy in the so-called radium bombs of the time Radio-therapists now use radiocobalt units a thousand times as strong using the cobalt 60 produced in nuclear reactors

We can now produce in this country about 10 000 curies a year, and, with our new reactor DIDO in operation, we shall

PLATE I

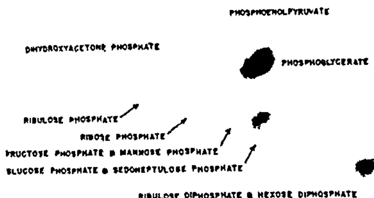


FIG 1 Chromatogram of extract from algae indicating uptake of radiocarbon during photosynthesis (10 seconds)

GLYCOLIC
GLYCERIC
MALIC
ALANINE
GLYCINE
SERINE
ASPARTIC
P-PYRUVIC
P-GLYCOLIC
PGA
TRIOSE P
P PENTOSE
MANNOSE P + FRUCTOSE P
SEDOHEPTULOSE P
GLUCOSE-P
SCENEDESMUS
60 Sec PS
RIBULOSE DI P
URIDINE DI P GLUCOSE

FIG 2 Chromatogram of extract from algae indicating uptake of radio carbon during photosynthesis (60 seconds)

and change of concept was leading to important revisions in the search for therapy of metabolic diseases However, only a beginning has been made—mainly by describing the metabolic processes—we are apparently still far from understanding in detail how they work It remains for the future to reveal how the new knowledge can be used for the conquest of disease'

In plant physiology the use of carbon 14 as a tracer has been of great value in the study of the mechanism of photosynthesis—the process by which plants take in carbon dioxide together with water to produce carbohydrates and oxygen The experiment consists of adding carbon dioxide containing carbon 14 to unlabelled carbon dioxide and this is absorbed by algae After quite a short time—10 seconds or 60 seconds for example—the plant is killed by sudden treatment with boiling ethanol, so that all enzymatic processes are halted Extracts of the plant material are made and two dimensional chromatography is used to study the synthesis of molecules A drop of the solution is put on to the corner of a large filter paper The solution then diffuses down and the different compounds are separated The filter paper is then turned through a right angle and the compound again diffuses down so that a two dimensional separation is achieved A photographic film is then placed against the filter paper and any molecules which contain carbon 14 blacken the film as in Plate I, Figures 1 and 2 [1] After 10 seconds phosphoglyceric acid is predominant so that it must be the first compound into which carbon dioxide is built By the end of 60 seconds sugar phosphates and acid phosphates have appeared The process of build up of the carbohydrates, proteins and fats is an extremely complex one and a great deal of light has been thrown on the successive steps in the reaction

Radioactive isotopes are also used as sources of radiation for the treatment of malignant disease Twenty years ago radium was used in quantities of several grams for radio-therapy in the so called radium bombs' of the time Radio therapists now use radiocobalt units a thousand times as strong using the cobalt 60 produced in nuclear reactors

We can now produce in this country about 10 000 curies a year, and, with our new reactor DIDO in operation we shall

be able to increase our annual output by 30,000 curies a year We are also just beginning to provide 1,000 curie sources of radioactive caesium 137 extracted from radioactive waste products Radiocaesium has a half life of 27 years and emits a gamma ray of about half the energy of radiocobalt We hope that in three or four years time we shall be able to produce it at lower costs than radiocobalt Because of its long half life it will be suitable for doing the work of 200 kilovolt X ray sets as well as having important industrial applications

Besides their use as an alternative to radium and X ray sets, radioisotopes have a special value in therapy because of their flexibility in use Thus radioactive sodium which emits a gamma ray has been injected as a solution into the bladder to treat malignancy, colloidal solutions of radioactive gold 199 or radioactive yttrium are injected into lung tumours because they stay reasonably well located and also decay rapidly with a half life of a few days

The use of thulium 170 and xenon 133 has been developed by Professor Mayneord and his colleagues These sources emit gamma rays of low energy—83 kV and 81 kV So the sources are easily protected and can be used in localities where X rays would be inapplicable For example it is possible to take pictures from inside the rectum showing the lower end of the vertebral column, or from inside the mouth showing one side of the jaw (Plate II, Figure 3)

A new method of treating brain tumours is being tried at Brookhaven in the United States Compounds containing boron 10 are injected intravenously into the patient Rapid permeation of the tumour occurs so that a higher concentration builds up there than in the brain The tumour is then irradiated with a neutron beam from a reactor for a period between 10 and 50 minutes after injection The neutrons disintegrate the boron 10 and a high speed α particle is emitted which affects the tumour The present statistics show a significant lengthening of survival time of the patients

An interesting new development is the application of radiation to viruses to change their structure and eliminate their virulence but leave their antigenicity unaffected Some interesting



FIG 3 Radiograph of dried skull taken with weak source of thulium 170 (from Mayneord W V (1952) *Brit J Radiol* 25 517)

insect pests and weevils is 25 to 50 per cent, other foodstuffs also suffer great losses Radiation has been used experimentally to destroy insect infestation in cereals and grain products at costs which are favourable compared with chemical methods The sprouting of potatoes can be inhibited by the effect of radiation on the enzyme system (Plate IV, Figure 5) The shelf life of certain foodstuffs can be substantially increased The economic importance of this can be shown by the fact that one British manufacturer lost £1 million in the hot summer of 1955 by deterioration of his products

We are in the process of founding the Wantage Radiation Laboratory where large sources of radiation will be available to study food preservation and other problems requiring high radiation rates

Radiation Hazards and Protection

This widespread use of radiation in hospitals research laboratories and industry will undoubtedly increase greatly and it is important to realize that radiation must be used with care and in conformity with strict codes of practice This has been made clear by a recent report by the Medical Research Council on the hazards to man of nuclear and allied radiations

Radiation produces mutations in human germ cells of a kind exactly similar to the natural mutations The effect of the mutations varies widely in severity Thus about one in 500 children born today suffers from a severe mental deficiency If the number of mutations were to be doubled by radiation there would be an increase of 1 500 cases of this particular deficiency in the next generation The number of hospital beds required, due to additional schizophrenia and manic depressive cases would increase by 400 The number of cases requiring medical care would be 5 to 10 times larger This is only one example of the effects of doubling mutation rates

In addition to these major ills resulting from mutations there would be an accumulation of less important mutations which would increase the general load of ill health in the population

For all these reasons geneticists are unanimous in saying that a doubling of the mutation rate of the whole population could

results have been obtained with a strain of foot and mouth disease virus There is a possibility here of a new method of producing vaccines

Radiation has proved to have a role in plant breeding by enabling the natural mutation rate to be increased a hundred fold The advantage of radiation is that the plant breeder does not have to wait so long before finding a mutant with desirable properties Examples of the way mutations affect a strain of barley are shown in Plate III Figure 4 [1] In the United States this method has produced a rust resisting strain of oats which is of economic importance In Canada a new strain of barley has been produced which matures earlier and so enables barley to be grown successfully in regions further north where the summer is short

Radiation has also been used successfully to control some insect pests Thus the screw worm has been almost eliminated in some areas of North America by sterilizing male flies by doses of radiation The male flies are then released, and since the female mates only once the production of new flies is greatly reduced

The radiation of hog carcases has also made possible the control of the parasitic disease Trichinosis which is transmitted by larvae in pork The irradiation sterilizes the female and breaks the disease cycle

In the country the Forest Products Research Laboratory is interested in the use of radiation for the preservation of timber of historic buildings from attacks of deathwatch beetle and other insect pests Research has shown that the fertility of beetles can be greatly reduced by doses of a few thousand roentgens This work has not yet reached the stage of applications to particular buildings

In agricultural research the use of radioisotopes as tracers is of value in the study of the comparative utilization of fertilizers by crops, in the method of action of insecticides and selective weedkillers and many other problems

Very active work is proceeding on the preservation of food stuffs In tropical parts of the world the loss of grain cereals by

[1] Plates III-V will be found between pp 12-13

should be reviewed to reduce this dose There is little doubt that substantial reductions can be made

Figure 6 shows the relative magnitude of different sources of radiation The contribution to the total population dose due to

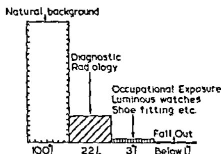


FIG 6 A comparison of various radiological doses

exposure of industrial workers to radiation sources in industry is about 3 per cent Codes of practice and regulation are soon to be issued to bring this more under control The contribution to the total population dose from the exposures of 20,000 atomic energy workers is less than one tenth of one per cent

Besides protecting the general population we have to protect individual workers and here again the M R C report provides valuable guidance In the past individual workers have been protected by a requirement to limit the dose of radiation to 0 3 roentgens per week which is at most 15 roentgens per year The M R C now make two additional recommendations First the dose to the age of 30 should not exceed 50 roentgens The reason for this is to keep the odds against producing harmful mutants in the individual s children to little less than the natural odds So the exposure of individuals should be reduced to a maximum of 5 roentgens a year

The second recommendation limits the lifetime dose to 200 roentgens This is a protection against radiation induced leukaemia and shortening of life span The natural incidence of leukaemia is at present one in 20,000 a year So in the Atomic Energy Authority staff of about 20 000 we should expect one

not be accepted, and Dr Muller, one of the most eminent of geneticists, has said that an increase of 25 per cent would be the most which could be tolerated

One important and difficult question is what radiation dose is required to double the mutation rate? Experimental evidence on human beings is negligibly small Experiments have been carried out mainly on fruit flies and mice These experiments led the Medical Research Council committee to conclude that the best estimate they can make in the light of present know ledge is that the doubling dose lies between 30 and 80 roentgens

Their overall assessment of the tolerable dose to the whole population has been given in terms of the radiation to which we are naturally subjected by the cosmic rays showered down upon us from outer space, from radiopotassium in our food and from radium in the ground and walls of our houses From these we can receive in our first 30 years of life (which are most important genetically) a dose to the gonads of about 3 roent gens

The M R C committee consider that it is unlikely that when the matter has been given further consideration an additional dose of more than twice the natural dose will be considered tolerable In other words the additional dose to the gonads due to all man made radiation should be less than 6 roentgens It seems possible that research into genetics and particularly human genetics will be greatly increased, in order to provide information about effects of radiation on mutation rates

The committee also collected figures for our present dose from man made radiations By far the greatest dose was from diagnostic radiology From a sample survey in some well run hospitals and from the number of radiological exposures now made in the United Kingdom, which amount to 12 million per annum the average dose is 0 66 roentgens in 30 years or 22 per cent of the natural dose The greater part of the exposure to the gonads is from examination of the pelvis and spine A similar survey made in the United States showed that the average dose was over four times larger than in the United Kingdom and reached 5 roentgens The M R C committee recommended that present practice in diagnostic radiology

the useful and long lived components such as radiocaesium and use these under safe and controlled conditions for medical and industrial purposes This will reduce the storage problem for the residual components

We must also guard against the possibility of reactors getting out of control, which would lead to melting of the fuel elements and the release of radioactive products Our first line of defence is to design reactors to have good inherent safety characteristics and to provide also a multiplicity of automatic safety devices which shut down the reactor if anything goes wrong This is combined with strict operating codes Our second line of defence is the containment of the reactor inside a closed pressure vessel The reactors of Calder Hall are so enclosed In the case of experimental reactors a third line of defence is added by containment of the whole reactor inside a steel sphere or cylinder The Dounreay and DIDO reactors are examples of this (Plate V, Figure 7)

The M R C report discusses the radiation dose due to fall out from the testing of atomic weapons At present the external radiation dose is small and in this country contributes less than one tenth of 1 per cent of the natural dose when account is taken of shielding by houses and weathering of the deposited activity It could rise to 1 per cent in a hundred years if present rate of testing continues

A more important point is the amount of strontium 90 finding its way into human bones This comes by way of the food chain, starting with the grazing of contaminated grass by the cow The strontium 90 in the grass finds its way into milk and then finds its way into bones The amount of radiostrontium in human bones at the present time is measurable but small A new unit has been coined for this known as a strontium unit' This is one thousandth of the level considered to be tolerable for occupationally exposed workers The M R C report set a maximum level for the general population at 100 strontium units but was of the opinion that serious consideration would be necessary if the level rose appreciably above 10 strontium units At present the level in human bones in the United Kingdom is about one strontium unit, and three times higher in some parts of the world

natural case a year The effect of a life dose of 200 roentgens is expected, from the evidence of Dr Court Brown published in the M R C report, to increase the incidence about tenfold So if we had 200 staff exposed to 200 roentgens, we should expect to have only one additional leukaemia due to radiation every 10 years Since the numbers exposed to these high doses are small, it is possible to programme their work to keep the dose to within the specified level

An often posed question is the future effect of generating a large part of our electricity from nuclear power, since it seems probable that by 1970 all new power stations in Britain will derive their heat from atomic energy

Our experience of operating existing nuclear reactors at Harwell and Windscale has shown that the exposure of reactor operators can be kept to below 5 roentgens per year, so that the hazards to individuals can be effectively prevented

The number of operators of nuclear power stations seems unlikely to exceed 10,000 by 1975 and, if their radiation dose is limited to 50 roentgens by age 30 the contribution to the average *population dose would be only one thousandth of a roent*gen or one three thousandth of the natural radiation dose Since all the chemical processing of spent fuel will be concentrated in one or two large plants, and since their operators will be limited to 5 roentgens per year their contribution to the population dose will also be negligible

The main danger we have to guard against is the release of large quantities of radioactive effluent from the chemical separation plants where spent fuel is processed either in the form of gaseous effluent or in liquid effluents or as a result of accidents

As a result of the work of the last decade by the Medical Research Council and the experimental work of a number of organizations, we have established codes of practice governing the discharge of radio-effluents and there should be no difficulty in continuing to adhere to safe practices The important point is that the highly radioactive wastes will be concentrated into a relatively small volume—a few cubic metres a year—and they can be safely stored in large tanks following our present practice We expect also to extract from the radioactive waste some of

PLATE III



Fi 4 Mutants with short straw

World wide measurements of radio strontium in foodstuffs have been undertaken by the United States Atomic Energy Commission We are carrying out similar measurements in this country and there is likely to be a great increase of such monitoring as a result of the formation of the United Nations Committee on Radiation

Radiological protection has the great advantage that radiation doses to individuals can be accurately and simply measured This is a much better position than in our exposure to viruses, carcinogens and other hazards of our present life We have also the advantage of a good tradition of radiological protection built up during the last ten years We believe therefore that in Britain atomic energy and its allied activities can be developed safely

The situation may not be the same in other countries which do not have the same experiences and traditions The United Nations has taken the lead by appointing an international committee to study these problems We hope that this will result in the adoption of universal codes of practice, similar to those in operation in Britain and the United States We hope also that if the International Atomic Energy Agency of the United Nations is established one of its principal cares will be to advise on radiation protection and to see that aid given through the Agency will be conditional on adherence to these codes of practice

The human race has a great deal to gain from this enormous potential increase in its source of energy and from the opportunity radiation provides to obtain a greater insight into living organisms Like all new activities of the human race these new powers introduce new risks We should not be afraid of these hazards, but should study and control them with the powerful tools of science and so enable the fruits of atomic energy to be enjoyed

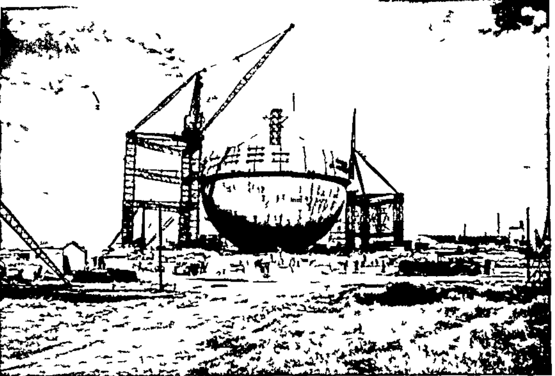


FIG 7 Sphere for Dounreay Fast Reactor in early stage of construction

PLATE IV

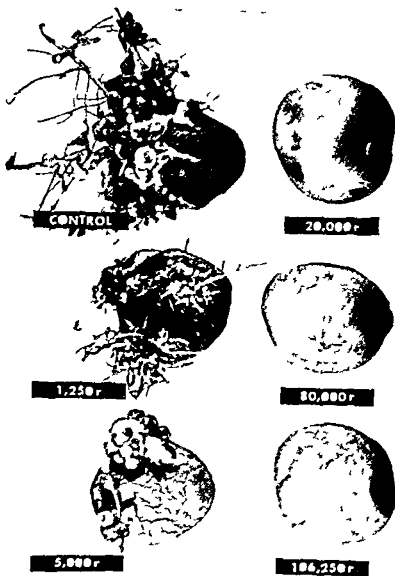


FIG 5 Prevention of sprouting in potatoes by irradiation at various doses

II

Homotransplantation of Organs

W J DEMPSTER

FROM a purely technical point of view, any organ belonging to the larger laboratory mammal can be transplanted provided that its blood vessels are of a reasonable calibre Whether such a transplanted organ will function or not is another matter An organ may be transplanted from its normal site to another site on the same animal and this procedure is referred to as autotransplantation An autotransplanted organ, such as the kidney, can function indefinitely A homotransplanted organ is one which is transferred from one animal to another of the same species A homotransplanted organ, such as the kidney may function for only a few days in dogs or for a few months in certain human recipients Organs such as the heart and lung are technically transplantable but function is uncertain even in the autotransplanted state amongst the factors complicating the transplantation of these organs which are still little understood are ventricular fibrillation and pulmonary oedema Transplantation of smaller organs such as thyroid ovary and adrenal, is more difficult One requires to transplant with these organs the large artery from which the organ distribution arises In the case of the thyroid it is the common carotid artery (Sterling and Goldsmith 1954), in the case of the ovary and adrenal it is the renal artery, and this involves transfer of the adjacent kidney with these organs (Levy and Blalock, 1939, Dempster, 1955b)

I wish to consider here the behaviour of the kidney, adrenal and ovary in the transplanted state I will make no further reference to the transplanted heart and lung because no sufficient

abnormal function appeared to be hydronephrosis which invariably develops to some extent in 'neck' kidneys Thus disturbance of function, however, required further investigation and, for this reason kidneys were autotransplanted to the pelvis The pelvic' kidney begins to function normally about three weeks after its transfer This immediate postoperative hyposthenuria is quite obscure, no doubt some enzyme processes have been temporarily disturbed Denervation and total temporary hyposthenuria do not appear to be involved since these procedures *per se*, do not affect markedly the concentrating power of a kidney In Table 1, the results of a series of experiments are listed It can be seen that neither denervation nor thirty minutes ischaemia seriously disturbs the concentrating capacity of a kidney

TABLE 1 Effects of denervation and ischaemia on urine output and specific gravity

Dog	S G before denervation	S G after denervation 2 hr samples	S G after ½ hr ischaemia	
			1st hr sample	successive 3 hrs sample
1	single kidney *L=1030	1030	1020	1028
2	R=1035 *L=1035	1035 1035	1035 1025	1035 1030
3	single kidney *L=1040	1040	1030	1035
4	R=1042 *L=1045	1043 1042	1042 1030	1042 1040
5	R=1037 *L=1035	1030 1035	1036 1022	1036 1030

* Denervated and rendered ischaemic

The fact that a kidney autotransplanted to the pelvis is unable to concentrate normally immediately after operation means that any definitive assessment of its function cannot be made until signs of recovery are evident It appears from the work of two groups of authors that the pelvic kidney usually regains its preoperative functional state within a month after transfer

data are available to allow any reasonable assessment to be made at present It is obviously a very difficult task to transplant hearts and lungs The kidney is easy to transplant and its function is easy to assess For this reason, I am in a position to review some of the laboratory and clinical work concerning kidney transplantation which has been reported in recent years

Transplantation means the transfer of an organ with an immediate re establishment of its circulation The results of this procedure have little in common with that in which fragments of tissue are transferred, this latter procedure is called implantation and, while frequently successful in rodents, is hardly ever of any value in larger laboratory mammals or in man

AUTOTRANSPLANTED KIDNEY

Before discussing the behaviour of the homotransplanted kidney, I wish to make some introductory remarks about the autotransplanted kidney Without adequate autotransplant control, one can become hopelessly confused by the occurrence of certain phenomena in homotransplanted kidneys

In the dog, the sites chosen for the reception of the kidney were the neck, in which situation the renal vessels were anastomosed to the carotido jugular circulation, and the pelvis, where the renal vessels were anastomosed to the iliac vessels In the former site, the ureter was brought on to the skin as a ureterostomy, but in the pelvis the ureter was reimplanted into the bladder The first technique has many advantages, one can keep the kidney under constant direct observation, one can note readily when it stops functioning one can remove biopsies or perform arteriogram —all without disturbing the animal very much All these advantages are of great importance in any study of the homotransplanted kidney In the human, homotransplantation of a kidney to the femoral region offers similar advantages (Hume Merrill, Miller and Thorn, 1955) One important disadvantage of the neck kidney is its inability to concentrate urine (Dempster and Joekes 1953) In spite of this a kidney can maintain life It can handle sodium and potassium very adequately but urea excretion is poor It is in a constant state of polyuria The most reasonable explanation for this

The natural history of homotransplanted kidneys follows a fairly well-defined pattern The homotransplanted kidney functions for a varying period of time and during this interval undergoes certain physical changes involving increase in size and weight, extensive details are available in the dog only (Dempster, 1955a) Part of the increased weight is due to considerable oedema of the renal sinus fat and perirenal tissues At some stage after a period of fairly adequate function, homotransplanted kidneys suddenly become anuric and at the same time the recipient becomes toxic (Dempster, 1953a) The cause of this type of anuria is obscure but it is thought to be the result of immunological reactions (Dempster, 1955a), this is the second type of anuria so far described The toxic syndrome is consistently severe in dogs (Dempster, 1953a) It has to be stressed that this clear cut toxic syndrome is not seen in bilaterally-nephrectomized dogs Dogs deprived of both kidneys sink into a ureamic state as the homotransplanted kidney begins to fail The toxic syndrome is clearly defined in *unilaterally nephrectomized dogs only* Thus, one *would not expect* this syndrome to appear in humans deprived of the function of both kidneys and in whom a homotransplanted kidney is failing as, indeed it did not appear in the experience of one group (Michon, Hamburger, Oeconomos Delinotte, Richet, Vaysse and Antoine, 1953) In chronic nephritics who are recipients of homotransplanted kidneys, it appears that a toxic syndrome can occur (Hume *et al* 1955) This group of workers describe the toxicity as mild compared to that reported as occurring in the dog It should be pointed out that some of these chronic nephritics received ACTH and cortisone Cortisone reduces the toxic state in dogs and indeed a pseudotoxic syndrome has been described (Dempster 1953c) A toxic state has been observed in goats but this does not appear to coincide with the imminent disintegration of the homotransplanted kidneys (Humphries 1956) In both dogs and in man, removal of the disintegrating kidneys brings prompt relief (Dempster, 1953b, Hume *et al* 1955) The cause of the toxic syndrome remains obscure (Dempster, 1953a)

If an arteriogram is taken of a homotransplanted kidney in

This sustained normal function has been followed in one dog by a group of workers for about two and a half years (Murray, Lang, Miller and Dammin, 1956) Another group studied several dogs for periods of up to nine months (Dempster, Joekes, and Oeconomos, 1955, Dempster, Eggleton and Shuster, 1956) In terms of human results, these follow up studies are trivial in length but they do give some indication of what to expect in the human

HOMOTRANSPLANTED KIDNEY

Usually, a homotransplanted kidney starts to function within a few minutes of establishing the new blood supply Sometimes, however, no urine appears or at the most a mere drop from time to time The same phenomenon occurs in autotransplanted kidneys and the cause is obscure (Dempster, 1954) In dogs, this anuria appears to be irreversible In humans, however, anurias of this type have been described but the kidneys have resumed function several weeks later (Hume *et al*, 1955) This type of anuria has been found to occur in goats (Humphries, 1956) The site of transfer may be significant since some authors (Egdahl and Hume, 1956), using exclusively the renal fossa site, have discounted the importance of this complication which they claim to have experienced on only one occasion Kidneys, whether auto- or homo-transplanted, which become immediately anuric after operation can be grouped into three histological types

1 those showing cloudy swelling of the tubular cytoplasm
2 those showing widespread casts
3 those showing widespread casts and extensive proximal tubular necrosis

The last type of lesion has been reported as occurring in a human homotransplanted kidney (Porter Joekes and Dempster, 1957) This type of anuria must be recognized since it could be confused with the genuine anurias associated with obscure immunological reactions (Dempster 1954) Indeed, in view of the experience of Hume *et al* 1955 it is imperative that one allows the immediately anuric kidney to remain undisturbed as function may be regained in a few days time

renal reticulum cells has been traced (Darmady, Dempster and Stranack, 1955) This cellular infiltration has been reported in dogs (Simonsen 1953a, Dempster 1953b) in humans (Michon *et al* 1953), and in goats (Humphries, 1956) It appears, then that this is a fairly generalized reaction of kidneys in the homotransplanted state If a nephron dissection is made, a curious lesion involving the neck of the tubule can be seen (Darmady *et al*, 1955) The area affected is collapsed and the epithelium appears translucent the remainder of the nephron appears normal This lesion can be prevented by administering cortisone Another feature shown up by the pyronine stains is the pyroninophilia of the renal vascular endothelium

2 If a kidney is allowed to remain in its host for about twenty four hours after the onset of anuria the histological picture is that of commencing disintegration Many tubules are necrotic and are surrounded by a cellular infiltration which is no longer composed only of transitional cells and immature plasma cells, polymorphs are now very much in evidence The glomerular tufts remain essentially normal This histological picture could quite easily be labelled—infection!

3 If a kidney is allowed to remain in its host for about forty eight hours after the onset of anuria, the histological picture is that of very advanced disintegration The tubules have largely melted away in an ocean of polymorphs Here and there one can detect the remains of a tuft which still appears to be normal The nature of this rapid lysis is still obscure

This somewhat artificial staging of the parenchymal disintegration has been worked out in detail in dogs only It would appear that the human material has been examined at stages 2 and 3 (Hume *et al*, 1955) It is of some considerable importance to be aware of this evolution of parenchymal disintegration In trying to assess the cause of anuria in homotransplanted kidneys one must consider the evidence presented in stage 1, and when this is done, it is very difficult to come to any conclusion If on the other hand, one examines stage 2 or 3 unaware of stage 1, it is easy to be misled into assuming that infection, if not altogether responsible for the anuria at least has to be taken into account

its anuric phase, an intense vasoconstriction is evident (Dempster, 1955a) The cause of this vasoconstriction has not been determined It was thought, at one time (Dempster, 1953b), that vasoconstriction, *per se*, was the cause of functional arrest of homotransplanted kidneys Experiments involving cortisone did not support this thesis (Dempster, 1953c) It would appear that the vasoconstriction is precipitated by factors other than mere stretching of the arterial tree induced by the enlarging kidney (Dempster, 1955a)

When a homotransplanted kidney stops functioning it usually *does so quite suddenly Certain qualifications require to be* added In dogs (Dempster, 1954) and in one normal human (Michon *et al*, 1953) the mode of sudden onset of anuria was identical In goats there is a slow decline In dogs on cortisone (Dempster, 1953c) there may be a slow decline in urine output In humans suffering from chronic nephritis (Hume *et al*, 1955) there would appear to be quite a different process at work Homotransplanted kidneys in such patients survive longer and the anuria is slow to establish itself (see also Murray and Holden (1954) who report survival for eighteen months) That the reaction of the chronic nephritic is somewhat different from the normal is naturally of great interest

When a homotransplanted kidney suddenly stops secreting, it is tempting to consider leaving well alone in case secretion returns The clinician, especially, is tempted to regard the anuria as the onset of a reversible tubular necrosis But it is not a wise policy to delay removing an anuric homotransplanted kidney because parenchymal disintegration ensues quite rapidly Three main stages in the histological appearances of anuric homotransplanted kidneys can be defined but it should be realized that each stage slips progressively into the other

1 If a kidney is removed within a few hours of becoming anuric, remarkably little tubular damage can be detected by the usual histological techniques One is attracted to the interstitial tissues by a marked infiltration of round cells, with pyronine stains (Dempster 1953b, Simonsen 1953a), these cells can be identified as immature plasma cells and the cells of Fagraeus (Fagraeus 1948) The evolution of these cells from

1 Infection This has been considered a likely cause in the past (Williamson, 1926) There is no evidence, from modern work, which would suggest that infection plays any important role The histological picture rules out this possibility provided one removes the homotransplanted kidney as soon as it becomes anuric

2 Increased interstitial pressure caused by oedema This is unlikely to be important since the tubules show no sign of collapse, on the contrary, many of the tubules appear dilated This of course can be explained, in part, by the mild hydronephrosis which develops in the early stages after a kidney is transplanted to the neck region of the dog

3 Hydronephrosis This is unlikely to be a cause of anuria in homotransplanted kidneys since this complication can occur in autotransplanted kidneys without anuria ensuing Together with this one can consider the possibility of vascular stretch. The kidney can tolerate marked degrees of vascular stretch without anuria ensuing (Dempster, 1955a)

4 Localized tubular lesion This lesion is abolished by treating the dogs with cortisone A lesion similar to this occurs in amyloid disease of the kidney without anuria ensuing Since cortisone abolishes the lesion but does not prolong the survival of homotransplanted kidneys, the tubular lesion would not appear to be the cause of the anuria

5 The immature plasma cell infiltration This is not likely to be the cause of anuria since plasma cell infiltrations are of everyday occurrence in chronically infected kidneys However both cortisone administration and local x irradiation can abolish the appearance of the immature plasma cells in homotransplanted kidneys (Dempster, 1953c) without any effect on the survival of such kidneys

6 Vascular spasm Evidence presented elsewhere (Dempster, 1955 see Figure 5 *a b*, *c*) would, at its face value suggest that vascular spasm was the cause of anuria However, cortisone can maintain a remarkably good flow through homotransplanted kidneys during their anuric phase, the vascular spasm is reduced and some patchy cortical filling occurs This reduction in no way prevents the onset of anuria Vascular spasm is, nonetheless,

It is understandable that in the clinic one hopes for an eventual resumption of function The advantage of using dogs, in this context, is that one can remove homotransplanted kidneys at any stage thought desirable In practice, however, the important reason for removing homotransplanted kidneys as soon as possible after the onset of anuria is on account of the toxic syndrome It subsides soon after the removal of the kidney

The next histological item one requires to consider is the capsule of a homotransplanted kidney This is markedly thickened and infiltrated with pyronine staining cells—such as have been described previously In certain instances, the kidney capsule becomes strongly adherent to the tissues of the host One should perform an intracapsular removal of such kidneys because important data can be obtained by a study of the capsule adherent to host tissue There is no evidence of a local cellular reaction (Dempster 1953b) Nor is there any evidence of a local cellular reaction around the homotransplanted adrenal and ovary Similar histological findings as regards immature plasma cells and the absence of a local reaction around the capsule have been reported in the homotransplanted adrenal and to some extent in the ovary (Dempster, 1955b) However, the ovary, at a time when the kidney and the adrenal appear histologically damaged, looks remarkably normal There is some evidence that homoimplanted ovaries are more able to survive their transfer to a foreign environment than most other tissues (Billingham and Parkes, 1955)

The local cellular reaction, so ably studied by Loeb (1947), has influenced studies on implantation of tissue for fifty years So far as one can detect at the moment, the local cellular reaction is quite irrelevant for tissues in the homotransplanted state, what its significance may be in the homoimplanted state is still not clear

Functional Arrest of Homotransplanted Kidneys

Why, after an interval of quite adequate function, does a homotransplanted kidney become anuric at some period after its transfer? This is by no means clear There are various possibilities which can be considered

1 Infection This has been considered a likely cause in the past (Williamson 1926) There is no evidence, from modern work, which would suggest that infection plays any important role The histological picture rules out this possibility provided one removes the homotransplanted kidney as soon as it becomes anuric

2 Increased interstitial pressure caused by oedema This is unlikely to be important since the tubules show no sign of col lapse, on the contrary, many of the tubules appear dilated This, of course, can be explained, in part, by the mild hydronephrosis which develops in the early stages after a kidney is transplanted to the neck region of the dog

3 Hydronephrosis This is unlikely to be a cause of anuria in homotransplanted kidneys since this complication can occur in autotransplanted kidneys without anuria ensuing Together with this one can consider the possibility of vascular stretch The kidney can tolerate marked degrees of vascular stretch without anuria ensuing (Dempster 1955a)

4 Localized tubular lesion This lesion is abolished by treating the dogs with cortisone A lesion similar to this occurs in amyloid disease of the kidney without anuria ensuing Since cortisone abolishes the lesion but does not prolong the survival of homotransplanted kidneys the tubular lesion would not appear to be the cause of the anuria

5 The immature plasma cell infiltration This is not likely to be the cause of anuria since plasma cell infiltrations are of everyday occurrence in chronically infected kidneys However both cortisone administration and local x irradiation can abolish the appearance of the immature plasma cells in homotransplanted kidneys (Dempster, 1953c) without any effect on the survival of such kidneys

6 Vascular spasm Evidence presented elsewhere (Dempster 1955 see Figure 5 *a* *b*, *c*) would, at its face value, suggest that vascular spasm was the cause of anuria However cortisone can maintain a remarkably good flow through homotransplanted kidneys during their anuric phase the vascular spasm is reduced and some patchy cortical filling occurs This reduction in no way prevents the onset of anuria Vascular spasm is, nonetheless

It is understandable that in the clinic one hopes for an eventual resumption of function The advantage of using dogs, in this context, is that one can remove homotransplanted kidneys at any stage thought desirable In practice, however, the important reason for removing homotransplanted kidneys as soon as possible after the onset of anuria is on account of the toxic syndrome It subsides soon after the removal of the kidney

The next histological item one requires to consider is the capsule of a homotransplanted kidney This is markedly thickened and infiltrated with pyronine staining cells—such as have been described previously In certain instances, the kidney capsule becomes strongly adherent to the tissues of the host One should perform an intracapsular removal of such kidneys because important data can be obtained by a study of the capsule adherent to host tissue There is no evidence of a local cellular reaction (Dempster 1953b) Nor is there any evidence of a local cellular reaction around the homotransplanted adrenal and ovary Similar histological findings as regards immature plasma cells and the absence of a local reaction around the capsule have been reported in the homotransplanted adrenal and to some extent in the ovary (Dempster, 1955b) However, the ovary, at a time when the kidney and the adrenal appear histologically damaged, looks remarkably normal There is some evidence that homoimplanted ovaries are more able to survive their transfer to a foreign environment than most other tissues (Billingham and Parkes 1955)

The local cellular reaction, so ably studied by Loeb (1947), has influenced studies on implantation of tissue for fifty years So far as one can detect at the moment the local cellular reaction is quite irrelevant for tissues in the homotransplanted state, what its significance may be in the homoimplanted state is still not clear

Functional Arrest of Homotransplanted Kidneys

Why, after an interval of quite adequate function, does a homotransplanted kidney become anuric at some period after its transfer? This is by no means clear There are various possibilities which can be considered

can be seen as a precipitate in the subcapsular space A whole range of intermediate reactions can and have been produced experimentally

There is considerable evidence that the fibrinoid material which one can see in the media of vessels in cases of periarteritis nodosa is a manifestation of an antigen antibody reaction A similar lesion occurs in scleroderma and other obscure collagen diseases In all these diseases the renal manifestation takes the form of a fibrinoid necrosis of the media of the blood vessels and the tuft capillaries Other than these two main groups of lesions we know nothing of how a kidney manifests an antigen-antibody reaction

If we approach the problem of what constitutes evidence of an antigen antibody reaction in a homotransplanted kidney by keeping in mind that any given tissue can react to trauma in a limited number of ways only certain points emerge for discussion

No evidence whatsoever of an antigen antibody reaction can be found in an anuric homotransplanted kidney—of a human a goat or a dog Most workers agree that the tufts remain essentially normal when all the surrounding tissue is disintegrating (Simonsen 1953a, Dempster, 1953b Hume *et al*, 1955, Humphries, 1956)

One might argue however, since *hetero* antigens and *hetero* antibodies are used in experimental nephritis and *hetero* antigens in the form of bacteria, may be involved in the clinical forms of glomerulo-nephritis—that our preceding criteria are not valid for *homo antigen antibody* reactions This is a perfectly permissible objection However if one analyses the histology of a second homotransplanted kidney one sees all the signs of the established criteria listed above (Dempster 1953b) This is unequivocal evidence that when an antigen antibody reaction does occur in a homotransplanted kidney, the nature of the tissue damage and the site of reaction correspond to what could be predicted from clinical and experimental evidence The above objection therefore, is not valid so far as one can judge at the moment

The evidence so far presented would support the theory of an actively acquired immunity on the basis of the accelerated

closely associated with whatever happens to the kidney in the immediate pre anuric phase

7 Immunological processes These are by no means clear and a consideration of this aspect of the problem will now be presented

IMMUNOLOGICAL REACTIONS

We must pass now to consider any possible immunological cause for the arrest of function of homotransplanted kidneys The theory of an actively acquired immunity, on the part of a host, is based on the accelerated disintegration of a second set of skin grafts on a recipient which had already received skin from the same donor (Medawar, 1944) From this sort of data, it has also been inferred that the first skin graft evoked and was destroyed by host antibodies The theory has been elaborated over the years and has been combined into brilliant work concerning the production of an actively acquired tolerance to skin grafts (Billingham, Brent, and Medawar, 1956)

A fair amount is known about the tissue damage concerned in renal antigen antibody reactions and this will now be considered There has been for a long time some indirect evidence that an antigen antibody reaction is involved in such diseases as glomerulo nephritis, renal amyloidosis and periarteritis nodosa The nature of these reactions was obscure Recent evidence, which is of a more positive kind, has shown that gamma globulins can be demonstrated in all these natural and experimental renal lesions (Mellors and Ortega, 1956) Globulins have been demonstrated, by the Coons test (Coons and Kaplan 1950), in the glomeruli in glomerulo nephritis and in the medial coat in periarteritis nodosa

The glomerular lesions produced in experimental nephritis were considered to be a manifestation of an an antigen antibody reaction—the nature of which was obscure Various interpretations have been put forward to explain the evolution of these lesions (Simonsen 1953b, Lippman 1955) Most workers are agreed that a mild reaction produces endothelial proliferation of the tuft and that a severe reaction produces wholesale renal damage in the course of which the tuft vessels are usually torn asunder or leak large quantities of protein which

therefore, should be unable to produce antibody or, if antibody is produced, it should be of low avidity Histological evidence does not support the latter possibility If one states quite frankly that the cause of arrest of function of first homotransplanted kidneys is not obvious it leaves the problem open for further consideration When a kidney is removed after the onset of anuria, the further elimination of antigen from that kidney ceases and all traces in the blood stream are rapidly swept into the reticulo endothelial system the host is now in a position to produce circulating antibody By the time a second kidney is transplanted, antibodies have been manufactured and are deployed against this kidney with resulting tissue damage typical of a severe antigen antibody reaction The exact sequence of events remains obscure The second kidney secretes soon after the establishment of its new circulation Some hours later anuria sets in and by twelve hours the kidney is totally destroyed Are antibodies circulating at the time of the operation? Is kidney antigen required to reach the reticulo endothelial system of the host in order to evoke this antibody response? There is no definite answer at the moment This very rapid destruction of a second kidney is at variance with the relatively slow destruction of a second skin homograft, six days is the average figure (Medawar, 1944) It may be that the assessment of the survival of skin homografts on the basis of histological integrity is not valid

Certain aspects of the histological features of homotransplanted kidneys remain to be assessed These relate to the significance of the immature plasma cells and the endothelial reaction observed in first kidneys There is a great deal of circumstantial evidence which suggests that the immature plasma cell has to do with antibody production (Fagraeus 1948) Can we conclude from this that the homotransplanted kidney is actually producing antibody against its host? The pyronine positivity of the vascular endothelial cells becomes of immunological significance if one accepts the opinion of Maximow (1924) that endothelial cells are merely flattened reticulum cells This activity of interstitial and vascular endothelial cells may represent a mobilization of the renal reticulo endothelial system to

disintegration of a second kidney, or of a first kidney following a skin graft (Dempster, 1953d) The cause of arrest of function of first kidneys remains obscure (Dempster, 1955) We can say that there is no sign of an antigen antibody reaction in a kidney homotransplanted to a non sensitized host One further piece of evidence must now be presented If a kidney be homotransplanted for four days (Simonsen, 1953a), or even for two days (Dempster, 1955a), and then returned to its original donor, it will continue to secrete for several days and then suddenly will become anuric The histological features seen in this kind of kidney resemble those seen in undisturbed homotransplants If a kidney is homotransplanted for twenty four hours and then returned to the original donor, it will continue to secrete indefinitely and, histologically, it remains normal All that can be concluded from this experiment is that some element of the host enters the interstitium of a homotransplanted kidney and starts a process which that kidney is quite unable to control even when it is returned to its normal environment Host antibodies cannot be claimed to have caused the anuria and the disintegration under these conditions The cause of the subsequent complete disintegration of the homo autotransplanted kidney remains obscure

Much circumstantial evidence indicates that host antibodies may not be involved against first homotransplanted kidneys This can be listed as follows

1 Histological criteria of renal antigen antibody reactions are not detectable or obvious (Dempster 1955a)

2 Total body irradiation of the host in no way prolongs the survival or in any way impedes the histological changes in homotransplanted kidneys (Dempster 1953b)

3 The embarrassment of antibody production so long as antigen remains in the blood stream should be considered Recent work has shown that the elimination of antigen from the blood stream precedes the appearance of antibody (Talmadge, Dixon, Bukantz and Dammin 1951) A homotransplanted kidney, placed as it is in the blood stream of the host, must eliminate antigen mainly via the renal vein One must assume that the elimination of antigen is continuous The host,

Dempster W J Joekes A M and Oeconomos N (1955) *Ann Roy Coll Surg Engl* **16** 324
Dempster W J Eggleton G M and Shuster S (1956) *J Physiol* **132** 213
Egdahl R and Hume, D M (1955) *Surg Gynec Obstet* **102** 450
Fagraeus A (1948) *Acta med scand* Suppl 204
Goodman M Greenspon S A and Krakower C A (1955) *J Immunol* **76** 96
Hill A. G S and Cruickshank B (1953) *Brit J exp Path* **34** ?7
Hume D M Merrill J P Miller B F and Thorn G W (1955) *J clin Invest* **34** 327
Humphries A. L (1956) Personal Communication
Levy S E and Blalock H (1939) *Ann Surg* **109** 84
Lippman R W (1955) *Proc nat Acad Sci* **41** 418
Loeb L (1947) *The biological basis of individuality* C C. Thomas Spring field Illinois
Maximow A. A (1924) *Physiol Rev* **4** 533
Medawar P B (1944) *J Anat Lond* **78** 176
Mellors R C and Ortega L G (1956) *Amer J Path* **32** 455
Merrill J P Murray J E Harrison J H and Guild W R (1956) *J Amer med Ass* **160** 277
Michon L Hamburger J Oeconomos N Delinotte P Richet G Vaysse J and Antoine B (1953) *Pr med* **70** 1419
Muirhead E E and Groves M (1955) *Arch Path* **59** 223
Murray G and Holden R (1954) *Amer J Surg* **87** 508
Murray J E Lang S Miller B F and Dammin G J (1956) *Surg Gynec Obstet* **103** 15
Porter K Joekes A M and Dempster W J (1957) *Brit J Surg* **44** 609
Simonsen M Buemann J Gammeltoft A Jensen F and Jorgensen K (1953a) *Acta path microbiol scand* **32** 36
Simonsen M (1953b) *Acta path microbiol scand* **32** 85
Sterling J A and Goldsmith R (1954) *Surgery* **35** 624
Talmadge D W Dixon F J Bukantz S C and Dammin G I (1951) *J Immunol* **67** ?43
Williamson C S (1926) *J Urol* **16** 231

deal with some antigenic element of the host There is some evidence that the homotransplanted kidney may be producing antibody against the host Muirhead and Groves (1955) have reported that the red cells of the host become Coombs positive several days after the transfer of a kidney This finding is of very great significance for it may mean that the gamma globulin coating the host red cells may be of renal origin The source of the raw material used by the homotransplanted kidney in its synthesis of gamma globulin must be the host This complicated aspect of tissue transfer cannot be predicted from work on tissue such as skin For this and many other reasons one cannot attempt, therefore to formulate general laws of tissue transfer immunity at this stage That the cause of the incompatibility of homotransplanted tissue is in some way related to the gene pattern was known many years ago (Loeb 1947) Recent work has shown that this is true since kidneys have been successfully exchanged between uniovular twins (Merrill, Murray, Harrison, and Guild, 1956) While this work takes us no further forward in the elucidation of the processes at work in the disintegration of homotransplanted organs, it does indicate that, within a limited clinical field life can be saved by homotransplanted kidneys

REFERENCES

BILLINGHAM R E BRENT L. and MEDAWAR P B (1956) *Phil Trans Roy Soc* B 239 357

BILLINGHAM R E and PARKES A S (1955) *Proc Roy Soc* B 143 550

COONS A H and KAPLAN M H (1950) *J exp Med* 91 1

DARMADY E M DEMPSTER W J and STRANACK F (19..) *J Path Bact* 70 225

DEMPSTER W J (1950) *Ann Roy Coll Surg Engl* 7 275

DEMPSTER W J (1953a) *Acta med scand* 144 360

DEMPSTER W J (1953b) *Brit J Surg* 42 540

DEMPSTER W J (1953c) *Arch int Pharmacodyn* 95 253

DEMPSTER W J (1953d) *Brit J plast Surg* 5 228

DEMPSTER W J (1954) *Acta med scand* 148 91

DEMPSTER W J (1955a) *Brit J Urol* 27 66

DEMPSTER W J (1955b) *Brit J Surg* 42 540

DEMPSTER W J and JOEKES A M (1953) *Acta med scand* 147 99

clot and planted in fresh medium Larger explants are grown in various types of *watch glass cultures* I myself generally use a simple moist chamber (Figure 1*c*, *d*), consisting of a watch glass enclosed by a Petri dish carpeted with wet cotton wool

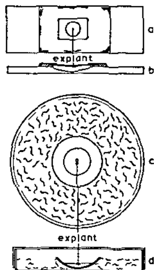


FIG 1 (*a*) Diagram of a hanging drop culture surface view (*b*) Same in longitudinal section (*c*) Diagram of a watch-glass culture surface view (*d*) Same in section (Fell 1940)

(Fell and Robison, 1929) The watch glass contains a clotted mixture of plasma and embryo extract and the explant grows on the surface of this medium Every two to three days the tissue is transplanted to a fresh clot I will mention a few examples of the various types of skeletal tissue that can be grown in a differentiated state by these simple techniques

(1) At one time it was sometimes thought that there was no such thing as an osteoblast and that bone was formed by fibroblasts in response to local humoral conditions In embryonic life this is certainly not true and the osteoblast represents a highly specialized type of cell which will form bone under the standardized humoral conditions of tissue culture (Fell, 1932)

III

The Physiology of Skeletal Tissue in Culture

HONOR B FELL

Introduction

THE fact that differentiated bone and cartilage can both be grown in tissue culture opens up many new possibilities for studying the physiology of skeletal tissues under simplified and controlled conditions

At first experiments were concerned mainly with the effect on skeletal explants of various chemical agents such as hormones and vitamins but recently the development of synthetic culture media and the synthesis of labelled metabolites have greatly extended the scope of physiological studies on skeletal cultures

I will first give you some idea of how skeletal tissues behave *in vitro* and of the different types of skeletal cultures that can be produced and then go on to tell you about some of the physiological observations that have been made on such material during the past few years Part of the work I am going to describe has not yet been published

Types of Skeletal Cultures

If you only want to grow small pieces of cartilage and bone, the large version of the old fashioned *hanging drop method* is quite satisfactory (Figure 1*a*, *b*) The piece of tissue is embedded in a clot composed of plasma and embryo extract on the surface of a coverslip The coverslip is then sealed over a hollow ground slide and incubated at body temperature Every two to three days the preparations are opened, and the tissue is cut out of the

Biochemical Studies

Phosphatase activity

The first biochemical study to be made on skeletal explants *in vitro* was an investigation of the *phosphatase activity* of femora from 6-day chick embryos, after different periods of cultivation

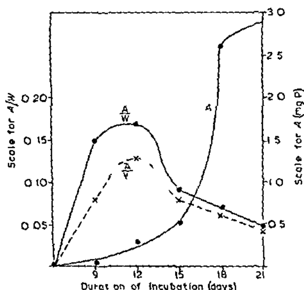


Fig 6 Production of phosphatase in the femur of the embryonic fowl during normal development *in vivo*

A = phosphatase per femur given as the amount of hydrolysis (mg P) of sodium glycerophosphate in 24 hrs at 37 C and pH 8.9

A/W = phosphatase per mg dry weight of femur (weighed after extraction)

A/W_1 = phosphatase per mg dry weight of femur (calculated on dry weight of corresponding unextracted femur)

(Fell and Robison 1929)

(Fell and Robison 1929) Robison found that normal hypertrophic cartilage from the intact chick embryo contains large amounts of phosphatase (Figure 6) We were able to show that this enzyme was also produced by the femora growing in culture (Figure 7)

If we examine a section of the normal tibia of a 6½ day chick embryo (Plate VI, Figure 2*a*)[1] we see that the cartilage cells in the middle segment of the shaft have begun to hypertrophy and in this region the perichondrium has differentiated into an inner layer of osteoblasts and an outer layer of fibroblasts If we strip the perichondrium from this middle region and grow it *in vitro* in a hanging drop preparation it will form a plate of bone (Plate VI, Figure 2*b*)

(2) In those parts of the body which will form both cartilage and membrane bone, the apparently undifferentiated mesoderm is already divided into sharply defined chondrogenic and osteogenic regions which are self differentiating when grown in culture The mandible of a 3–4 day chick embryo provides a good example of this regional specificity (Jacobson and Fell, 1941) When a block of tissue is excised from the mandible of a 3 day chick embryo (Plate VI, Figure 3*a*) and grown *in vitro*, a cartilaginous rod differentiates during cultivation (Plate VI Figure 3*b*), because the explant contained the chondrogenic region destined to form Meckel s cartilage Other areas, histologically indistinguishable from the rest of the mandible, are predetermined to form membrane bone Thus when part of the proximal region of the mandible of a 4 day embryo is cultivated (Plate VII, Figure 4*a*), an ossification centre differentiates (Plate VII, Figure 4*b*) because potential osteogenic tissue was included in the explant It is a curious and rather paradoxical fact that in the apparently undifferentiated mesoderm of the early embryo, the chondrogenic and osteogenic areas should be so rigidly determined, whereas in the adult animal ectopic bone and cartilage can develop almost anywhere in the body

(3) Entire skeletal rudiments do very well when grown by the watch glass method (Plate VII Figure 5), the rudiments enlarge to several times their original size and continue to develop both anatomically and histologically We are using such explants extensively for studies on the physiology of skeletal tissue *in vitro*

[1] The plates referred to in this lecture will be found between pages 36–7 and 44–5

Biochemical Studies

Phosphatase activity

The first biochemical study to be made on skeletal explants *in vitro* was an investigation of the *phosphatase activity* of femora from 6-day chick embryos after different periods of cultivation

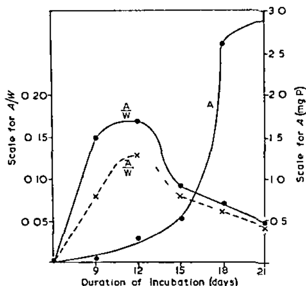


Fig 6 Production of phosphatase in the femur of the embryonic fowl during normal development *in vivo*

A = phosphatase per femur given as the amount of hydrolysis (mg P) of sodium glycerophosphate in 24 hrs at 37 C and pH 8 5

A/W = phosphatase per mg dry weight of femur (weighed after extraction)

A/W_1 = phosphatase per mg dry weight of femur (calculated on dry weight of corresponding unextracted femur)

(Fell and Robison 1929)

(Fell and Robison 1929) Robison found that normal hypertrophic cartilage from the intact chick embryo contains large amounts of phosphatase (Figure 6) We were able to show that this enzyme was also produced by the femora growing in culture (Figure 7)

Bone rudiments in synthetic medium

Bone rudiments can be grown in a chemically defined medium This was first shown by Dr Etienne Wolff and his collaborators (Wolff, Haffen, Kieny and Wolff, 1953) who

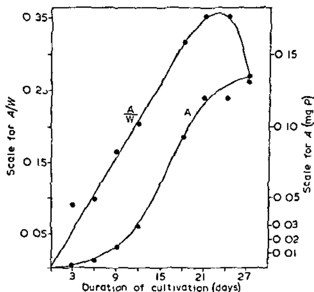


FIG 7 Production of phosphatase during the development *in vitro* of femora from 6-day fowl embryos
A = phosphatase per femur
A/W = phosphatase per mg dry weight of femur (weighed after extraction)

(Fell and Robison 1929)

used a very simple medium consisting of agar and a few amino acids Better growth (Figure 8) can be obtained in the media prepared by Dr Raymond Parker and his associates at Toronto University (Healy, Fisher and Parker, 1954 1955) The composition of these nutrient solutions is complex, the media most frequently u ed in our laboratory contain about sixty components which include some twenty amino acids as well as a range of vitamins coenzymes, nucleotides glucose and salts These chemically defined solutions in contrast to the natural media used hitherto enable the nutritional requirements and

metabolic activities of organ and cell cultures to be studied, and certain aspects of the metabolism of isolated bone rudiments in Parker s media Nos 858 and 929 have already been investigated by Dr Biggers in collaboration with Dr Webb Dr Parker and Dr Healy (1957)

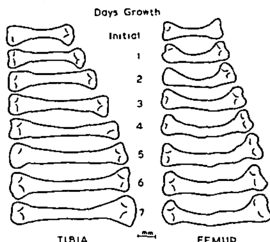


FIG 8 Camera lucida drawings of tibia and femur from a 7-day chick embryo during 7 days cultivation in Parker s synthetic medium 929 (Biggers Webb Parker and Healy 1957)

In these experiments the bones were cultivated by Chen's modification of the watch glass method (Chen 1954), in this technique the explants are grown on rafts of lens paper floating on the surface of the fluid medium During the culture period there is an increase in length (Figure 9) and also in the total nitrogen content of the explants, although this is less than in serum During cultivation considerable amounts of the amino acids are utilized This has been shown by estimations of the total amino nitrogen of the medium before and after cultivation (Figure 10), while the utilization of the individual amino acids of the medium has been followed by Moore and Stein s (1951) chromatographic technique (Figure 11) an illustration of the application of this method to tissue culture work has been given by Biggers, Rinaldini and Webb (1957)

Although the incorporation of an amino acid into the protein of the bones has been demonstrated by the inclusion of glycine-^{14}C in the medium (Biggers, Webb, Parker and Healy, 1957), considerable amino acid deamination must occur during the

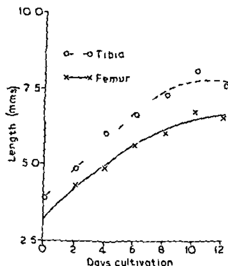


Fig 9 Average growth curves of tibia and femur in synthetic medium 929 (Biggers Webb Parker and Healy 1957)

culture period since only a fraction of the amino nitrogen which disappears from the medium is taken up into the explants

To find which amino acids of the defined media are essential for growth, Dr Biggers prepared twenty solutions, each deficient in one amino acid and compared increase in weight of bones cultivated on these deficient media and on a control medium containing all the amino acids His results showed that eleven amino acids were essential (Plate VIII, Figure 12) and nine non essential for normal growth The requirements for some of the essential amino acids (e g threonine, arginine and trypophan) were such that the effects of their deficiency were obvious after only 48 hours growth with others (e g lysine), *signs of* deficiency appeared more slowly

Although nine amino acids were judged non essential on the basis of the absence of any significant response to their omission

from the medium, they are present as essential components of the proteins of the rudiments In consequence, Dr Biggers and Dr Webb have investigated the possible synthesis of these amino acids from other components of the nutrient Glucose is

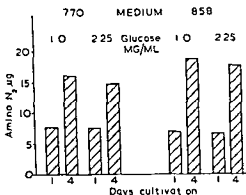


FIG 10 Histograms showing the total amino-nitrogen utilized (i e the percentage removed from the culture medium) by femora and tibiae from 7-day chick embryos grown in synthetic media supplemented with additional glucose (Biggers Webb Parker and Healy 1957)

the source of six of these This has been established by experiments in which bone rudiments were grown for 24 hours in a defined medium containing glucose generally labelled with ^{14}C After growth the bones were fractionated to isolate the protein which was then hydrolysed and the radio active amino acids located by autoradiography of the paper chromatograms were as follows alanine glutamic acid, aspartic acid serine glycine and proline

Biosynthesis of collagen in osteoblast cultures

Dr Fitton Jackson and Dr Ronald Smith (Fitton Jackson and Smith 1957 Smith and Fitton Jackson, 1957) have been using cultures of osteoblasts to study the biosynthesis of collagen The frontal bones of 11 day chick embryos were disintegrated into suspensions of isolated cells by tryptic digestion, the osteoblasts were then cultivated in a fluid, fibrin free medium on a sheet of glass by a method devised by Dr Fitton Jackson

Although the incorporation of an amino acid into the protein of the bones has been demonstrated by the inclusion of glycine-^{14}C in the medium (Biggers, Webb, Parker and Healy, 1957), considerable amino acid deamination must occur during the

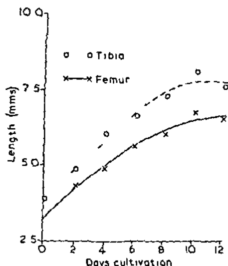


Fig 9 Average growth curves of tibia and femur in synthetic medium 929 (Biggers Webb Parker and Healy 1957)

culture period since only a fraction of the amino nitrogen which disappears from the medium is taken up into the explants

To find which amino acids of the defined media are essential for growth, Dr Biggers prepared twenty solutions each deficient in one amino acid and compared increase in weight of bones cultivated on these deficient media and on a control medium containing all the amino acids His results showed that eleven amino acids were essential (Plate VIII Figure 12) and nine non essential for normal growth The requirements for some of the essential amino acids (e g threonine arginine and trypophan) were such that the effects of their deficiency were obvious after only 48 hours growth with others (e g lysine), signs of deficiency appeared more slowly

Although nine amino acids were judged non essential on the basis of the absence of any significant response to their omission



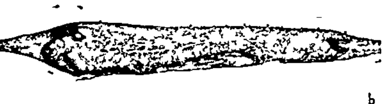


FIG 2 (*a*) Section of normal tibia from 6½ day chick embryo Control Note the region of hypertrophic cartilage cells in the middle segment of the shaft and the overlying two layered periosteum (*b*) Vertical section of the periosteum isolated from the opposite tibia of the same chick after 10 days cultivation in a hanging drop preparation A plate of bone has developed in culture (× 75)

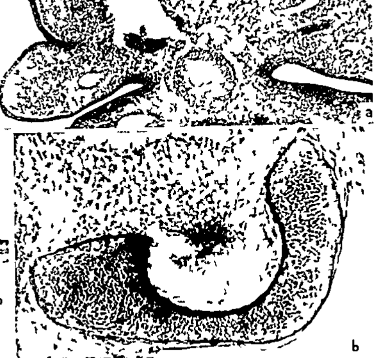


FIG 3 (*a*) Frontal section of the normal mandible of a 3 day chick embryo Control A block of tissue has been excised and cultivated *in vitro* (*b*) The fragment removed from the mandible seen in (*a*) after 8 days cultivation The explant contained the chondrogenic tissue and has differentiated into Meckel s cartilage (× 90)

(Jacobson and Fell 1941)

The cells spread out on the glass to form a thin sheet of tissue in which collagen fibrils appeared after about two days' growth (Plate VIII, Figure 13) and rapidly increased in number Cultures of this type provide a very convenient closed system in

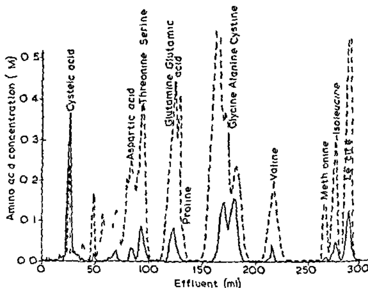


FIG 11 The utilization of individual amino acids analyses by Moore and Stein s chromatographic method showing the concentration of individual amino acids in unused synthetic medium No 858 (dotted line) and in the same medium after embryonic femora and tibiae had grown in it for 4 days (continuous line 4 bones per ml of medium) (Biggers Rinaldini and Webb 1957)

which the biochemical aspects of collagen formation can be studied in the absence of the metabolic influences of the body as a whole

First of all Fitton Jackson and Smith investigated the relationship between growth (as measured by increase in dry weight of the tissue) the formation of collagen protein (based on the estimation of hydroxyproline) and the appearance of typical collagen fibrils (as demonstrated by electron micrographs of pieces of tissue from the culture)

The results of these experiments showed that during the first 48 hours growth, there was a mean increase of 40–50 fold in

the amount of hydroxyproline present in the tissue (Figure 14), although, as I have just said, demonstrable collagen fibrils did not appear until about the second day or shortly afterwards Collagen fibres were formed profusely during

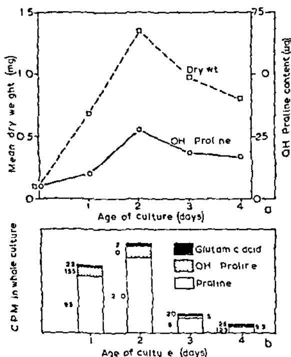


FIG 14 (*a*) Curves showing the increase in dry weight and formation of hydroxyproline (chemical estimation) in osteoblast cultures grown in the presence of ^{14}C proline and harvested at different stages (*b*) Histograms White total amount of labelled proline incorporated in the protein Stippled amount of protein bound ^{14}C-hydroxyproline Black amount of labelling in glutamic acid (Fitton Jackson and Smith)

the later stages of cultivation but the increasing production of fibrils was not accompanied by any significant rise in the mean hydroxyproline content of the tissue

Smith and Fitton Jackson then made experiments to see whether the osteoblasts in these cultures would convert proline into hydroxyproline (Figure 14) They added ^{14}C labelled

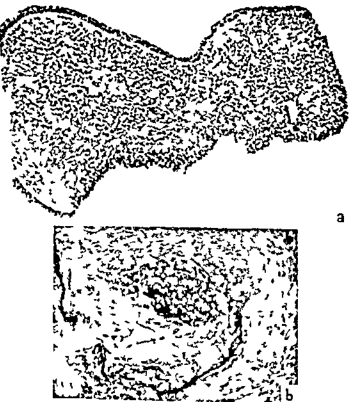


FIG 4 (*a*) Frontal section of the proximal region of the right half of the mandible of a 4-day chick embryo Control The dotted line indicates where a block of tissue has been removed and explanted *in vitro* (× 115) (*b*) The tissue removed from the mandible after 6 days cultivation The explant contained the osteogenic mesoderm and has formed an ossification centre (× 175)
([illegible] and Fell, [illegible])

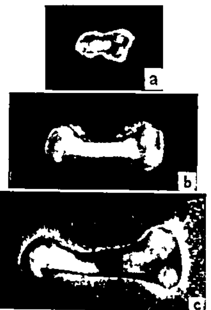


FIG 5 Serial photographs of the femur rudiment of a 5½ day chick embryo during growth in a watch glass culture (*a*) original explant (*b*) 7 days in culture (*c*) 14 days in culture
(From a film by Canti and Fell)

mouse embryos near term Though small, such bones are at an advanced stage of development, with a stout shaft of bone enclosing a narrow cavity, and large cartilaginous ends One of each pair of bones was grown in normal medium and the other in medium to which 1 500–3 000 i u of vitamin/100 ml of medium had been added

The controls remained healthy and intact for about ten days and enlarged slightly but the corresponding +A bones rapidly disappeared The cartilage matrix lost its metachromasia and shrunk and the bone was resorbed, although the surrounding soft tissue migrated into the medium and grew profusely (Plate IX Figures 15 and 16)

We then went on to study the effect of excess vitamin A on very early long bone rudiments from 6 day chick embryos In normal medium, as I have already shown, such rudiments grow to several times their original length, acquire a surprisingly normal shape, and continue their histological differentiation In the presence of excess vitamin A however, first growth was arrested and then the rudiment began to shrink and eventually disappeared almost completely although as in the case of the mouse bones, the soft tissue grew profusely (Figure 17) This shrinkage is accompanied by loss of the metachromatic staining of the matrix (Plate X Figure 18)

We were naturally anxious to find out something about the metabolic changes which accompanied these drastic histological effects At the time of Sir Edward s death, he and I were collaborating with Dr Stephen Pelc in studying the uptake of radio-active sulphate ($^{35}SO_4$) by normal and +A bone rudiments in culture Dr Pelc and I finished the investigation and the results have been published (Fell Mellanby and Pelc 1956) This work was done by means of autoradiography We used Dr Pelc s stripping film method in which photographic film is superimposed on histological sections (Doniach and Pelc 1950) After the autoradiograph has been developed, the underlying section can be suitably stained and the histology correlated with the autoradiograph

As before control bones were grown in normal medium and the +A bones in medium to which 2 800–3 000 i u /100 ml of

L-proline to the culture medium and studied its subsequent fate in the tissue A significant amount of the incorporated radio activity appeared in the hydroxyproline and glutamic acid fractions of the cultures The most rapid rate of formation of labelled hydroxyproline occurred during the first two days, i e before collagen fibres had appeared in the cultures, it then fell exponentially as the age of the cultures increased although the formation of collagen fibres had become very active From this and earlier work (Fitton Jackson, 1956), Fitton Jackson and Smith concluded that under the influence of the osteoblasts, proline is first incorporated in peptide linkage and part is then hydroxylated to form a hydroxyproline rich precursor which subsequently becomes directly transformed into typical collagen fibrils

THE EFFECT OF CHEMICAL AGENTS

At the beginning of this lecture I said that skeletal tissue in culture can be used to study the effects of physiologically active agents such as vitamins and hormones I will now give a few examples of the kind of results that can be obtained in such experiments

Vitamin A

Sir Edward Mellanby and I studied the effect of excess vitamin A on bone rudiments from embryonic mice and chicks (Fell and Mellanby, 1952) It has long been known that when young animals are fed on a diet containing excess vitamin A, there is a rapid resorption of cartilage, and rarefaction of the bone often causes spontaneous fractures But there was some doubt as to whether this effect was directly due to the vitamin or whether it was mediated through some other system or organ The object of our early experiments was to see whether vitamin A would have a direct action on skeletal rudiments isolated in culture, when it was added to the nutrient medium in con centrations similar to those found in the blood plasma of hypervitaminotic animals The bones were grown by the watch glass method which I have already described Our first experiments were made with the tibia, radius and ulna from

mouse embryos near term Though small, such bones are at an advanced stage of development with a stout shaft of bone enclosing a narrow cavity, and large cartilaginous ends One of each pair of bones was grown in normal medium and the other in medium to which 1,500–3 000 i u of vitamin/100 ml of medium had been added

The controls remained healthy and intact for about ten days and enlarged slightly but the corresponding +A bones rapidly disappeared The cartilage matrix lost its metachromasia and shrunk and the bone was resorbed, although the surrounding soft tissue migrated into the medium and grew profusely (Plate IX Figures 15 and 16)

We then went on to study the effect of excess vitamin A on very early long bone rudiments from 6 day chick embryos In normal medium as I have already shown such rudiments grow to several times their original length acquire a surprisingly normal shape and continue their histological differentiation In the presence of excess vitamin A, however first growth was arrested and then the rudiment began to shrink and eventually disappeared almost completely although, as in the case of the mouse bones the soft tissue grew profusely (Figure 17) This shrinkage is accompanied by loss of the metachromatic staining of the matrix (Plate X Figure 18)

We were naturally anxious to find out something about the metabolic changes which accompanied these drastic histological effects At the time of Sir Edward s death, he and I were collaborating with Dr Stephen Pelc in studying the uptake of radio-active sulphate ($^{35}SO_4$) by normal and +A bone rudiments in culture Dr Pelc and I finished the investigation and the results have been published (Fell Mellanby and Pelc 1956) This work was done by means of autoradiography We used Dr Pelc s stripping film method in which photographic film is superimposed on histological sections (Doniach and Pelc 1950) After the autoradiograph has been developed, the underlying section can be suitably stained and the histology correlated with the autoradiograph

As before control bones were grown in normal medium and the +A bones in medium to which 2 800–3 000 i u /100 ml of

vitamin A had been added One drop of Tyrode s solution containing 200 μc $Na_2{}^{35}SO_4$ was deposited on each explant and the excess was pipetted off after 30 seconds Some explants were fixed after 2, 4, 24 or 48 hours in labelled medium, others

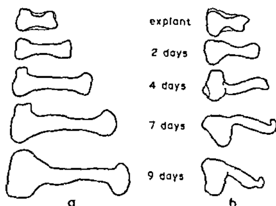


FIG 17 Camera lucida drawings of a pair of explanted femora from the same 6-day chick embryo (*a*) during 9 days cultivation in normal medium (*b*) during 9 days in medium containing excess vitamin A (1500 i u per 100 ml of medium) (Fell and Mellanby 1952)

were kept in labelled medium for 4 hours and fixed 24 or 48 hours after transfer to unlabelled medium

In the control explants treated with labelled sulphate after six days growth *in vitro*, the autoradiograph was heaviest over the newly formed hypertrophic cartilage and weakest over the oldest hypertrophic cartilage in the interior of the shaft, it was also dense over the epiphyses If a control bone was transferred from labelled to unlabelled medium for two days little bound sulphate was lost from the cartilage When +A explants were treated with $^{35}SO_4$ after four days growth they took up sulphate but if the bones were then transferred to unlabelled +A medium much of the sulphate was lost, especially from the middle segment of the shaft and the periphery of the rudiment After six days growth in +A medium the peripheral cartilage and the middle segment of the diaphysis ceased to incorporate sulphate This was correlated with the loss of metachromatic staining of the matrix with toluidine blue

From these results we concluded that excess vitamin A both inhibits the synthesis of new cartilage matrix and causes the dissolution of that already formed

Hormones

The effect of various hormones on skeletal tissue in culture has been studied

In our laboratory we have done a good deal of work on the action of thyroxine and related compounds on the development of long bone rudiments in culture This research was begun in collaboration with Sir Edward Mellanby and part of our recent work is being done in conjunction with Dr Rosalind Pitt Rivers and Mrs Valerie Galton

Thyroxine is known to affect the skeleton *in vivo* For example, Simpson, Asling and Evans (1950) found that injection of thyroxine into the young rat or mouse caused premature ageing of the skeleton, it accelerated the differentiation of the long bones but had no growth promoting effect So we had reason to believe that skeletal elements might prove suitable material for the investigations we had in mind The experiments I am going to describe had two main objects to see whether the compounds would act on the target organ under the simplified conditions of growth *in vitro* and to compare the relative potencies of the compounds if they proved to be active in culture

As experimental material we used the femur tibia humerus, ulna and radius of 6 day chick embryos The bones were dis sected from the chick and explanted in the usual way on the surface of a plasma embryo extract clot contained in a watch glass culture The effect of thyroxine and related compounds on such explants was investigated in the following way Of each pair of long bone rudiments one was grown in medium to which the compound had been added, and the other in normal medium as a control Every two days the bones were drawn with the aid of a camera lucida and their growth in length was measured from the serial drawings The compounds were added in concentrations similar to the levels of thyroxine found in the blood in cases of hyperthyroidism

The results showed that thyroxine (T_4) (Fell and Mellanby, 1955), triiodothyronine (T_3) (ibid, 1956) and triiodothyroacetic acid (TA_3) (Fell, Galton and Pitt Rivers, unpublished) all have a pronounced effect on the long bone rudiments in culture All three compounds accelerate differentiation Thus they hasten the formation of intercellular material in the epiphyses, so that the epiphyses are larger in the treated rudiments than in the controls, they also accelerate hypertrophy in the cells of the shaft which causes a great reduction in the width of the proliferative zones, because the cells hypertrophy faster than they are replaced by multiplication This effect is already seen after only four days in culture

An interesting fact came to light in these experiments for a given dose of hormone, the growth rates of the various long bone rudiments from the same chick were affected in widely different degrees (Plates X XI Figures 19, 20 and Figures 21, 22) For example the growth rate of the tibia and femur was always inhibited while that of the radius was stimulated This was true of all the three effective compounds so far investigated

One of my colleagues, Miss Kirstie Lawson has measured the wet weight and total nitrogen of the T_3 treated and control bones (personal communication) She finds that the T_3 treated tibiae are lighter and contain less nitrogen than the controls in normal medium, but the T_3 treated radii are heavier and contain more nitrogen than the corresponding controls

Why the various rudiments should show this constant difference in response is not entirely clear I investigated the normal differentiation of the five long bones in the embryo, and found that the rate of differentiation of the five rudiments, as indicated by the degree and extent of cellular hypertrophy in the shaft, varied in the same sense as their response to thyroxine and related compounds in culture The differentiation was most rapid in the tibia next in the femur then in the humerus, then in the ulna and slowest in the radius So the more rapidly the bones develop *in vivo* the more severely are they affected by the compounds *in vitro* But at present I do not know if there is any causal relationship between these two facts and Miss Lawson is investigating a number of other possibilities

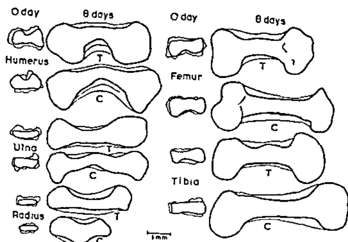


FIG 21 Camera lucida drawings of the living long bones from the legs and wings of a 6-day embryo One of each pair (T) was grown for 8 days in medium containing 16 µg thyroxine per 100 ml and the other (C) in control medium Note the differential effect of the hormone on the growth of the various rudiments (Fell and Mellanby 1955)

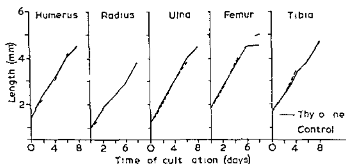


FIG 22 The differential effect of 16 µg/100 ml of thyroxine on the average growth rates of the humerus radius ulna femur and tibia of 6 day chick embryos during 8 days cultivation *in vitro* (Fell and Mellanby 1955)

We have done some experiments to compare the relative potencies of T4, T3 and TA3 To make these comparisons, bone rudiments from one side of each embryo were grown in medium containing one of these compounds, and those from the opposite side in medium containing another, the growth rates of the paired rudiments were then measured and plotted We found that in a given concentration, the inhibitory effect of T3 on the growth rates of the femora and tibiae was about four times that of T4

In the single experiment that we have done so far, T3 and TA3 seemed to be equally potent

Conclusion

In this lecture I have shown how bone and cartilage will develop in culture in a surprisingly normal way, how modern microchemical methods can be used to investigate the metabolism of these skeletal explants and how their structure can be modified by such agents as vitamins and hormones So far biochemical studies on skeletal tissue cultures have been concerned mainly with explants growing under as nearly normal conditions as possible, but we have just begun to make similar experiments on cultures of cartilage and bone treated with vitamins and hormones It is already obvious that such experiments are likely to extend our knowledge of the direct action of these agents on skeletal tissues and I expect that we shall see a considerable expansion of this field during the next few years

Acknowledgements

I am indebted to Drs J D Biggers, M Webb, S Fitton Jackson and Dr R H Smith for allowing me to mention their unpublished results and to Mr V C Norfield for preparing the plates I also wish to express my gratitude to the following for permission to reproduce illustrations the Society of Experimental Biology and the editors of the *Biochemical Journal* the *British Medical Journal* the *Journal of Physiology* the *Journal of the Royal Microscopical Society*, *Nature* and the *Quarterly Journal of Microscopical Sciences*

PLATE VIII



FIG 12 Living femur rudiments from the same 7 day embryo after 6 days cultivation in (*a*) normal synthetic medium No 858 and (*b*) the same medium but without arginine (Biggers Webb Parker and Healy 1957)

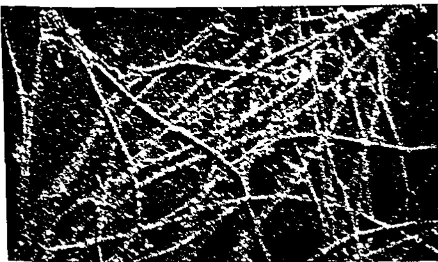


FIG 13 Electron micrograph of collagen fibrils formed in an osteoblast culture after 2 days growth (× 75 000)
(Fitton Jackson and Smith unpublished)

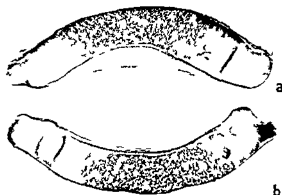


FIG 20 Ulnae from the same chick as the tibiae in Fig 10 after 6 days in culture (*a*) in normal medium (*b*) in medium containing 16 μg/100 ml of thyroxine Cellular hypertrophy is more advanced in (*b*) but there is no reduction in growth (× 20) (Fell and Mellanby 1955)

REFERENCES

BIGGERS J D RINALDINI L M and WEBB M (1957) *Symp Soc exp Biol* **11** 264

BIGGERS J D WEBB M PARKER R C and HEALY G M (1957) *Nature Lond* **180** 825

CHEN J M (1954) *Exp Cell Res* **7** 518

DONIACH I and PELC S R (1950) *Brit J Radiol* **23** 184

FELL H B (1932) *J Anat* **66** 157

FELL H B (1940) *J Roy mic Soc* **50**, 95

FELL H B and MELLANBY E (1950) *Brit med J* **11** 535

FELL H B and MELLANBY E (1952) *J Physiol* Lond **116** 320

FELL H B and MELLANBY E (1955) *J Physiol* Lond **127** 427

FELL H B and MELLANBY E (1956) *J Physiol* Lond **133** 89

FELL H B MELLANBY E and PELC S R (1956) *J Physiol* Lond **134** 179

FELL H B and ROBISON R (1929) *Biochem J* **23** 767

FITTON JACKSON S (1956) *Proc Roy Soc* B **144** 556

FITTON JACKSON S and SMITH R H (1957) *Biophys Biochem Cytol* **3** (in press)

HEALY G M FISHER D C and PARKER R C (1954) *Canad J Biochem Physiol* **32** 327

HEALY G M, FISHER D C and PARKER R C (1955) *Proc Soc exp Biol N Y* **89** 71

JACOBSON W and FELL H B (1941) *Quart J microscop Sci* **82** 563

MOORE S and STEIN W H (1951) *J biol Chem* **192** 663

SIMPSON M E ASLING C W and EVANS H M (1950) *Yale J Biol Med* **23** 1

SMITH R H and FITTON JACKSON S (1957) *Biophys Biochem Cytol* **3** (in press)

WOLFF E HAFFEN R KIENY M and WOLFF E (1953) *J Embryol exp Morphol* **1** 55

IV

Arterial Substitutes

CHARLES ROB

THE efforts of surgeons to repair blood vessels go back a fair way The first possibly successful operation of this type was carried out by Hallowell of Newcastle upon Tyne in 1759 at the suggestion of Lambert, and reported by him in 1764 He closed a lateral defect in the brachial artery and at no time did the patient lose his radial pulse Over 100 years were now to elapse before Jassinowsky (1889) using a careful aseptic technique succeeded in twenty two out of twenty six end to end anastomoses of the carotid arteries of sheep The modern technique of vascular anastomosis using two or three interrupted stay sutures and then a continuous everting mattress or over and over suture of fine silk was worked out by *Jaboulay and Brian (1896) and Carrel (1907) and has not altered* since The first successful end to-end arterial anastomosis in man was performed by Murphy (1897) using an invagination technique, and the first porta caval anastomosis in a patient was performed by Vidal in 1903 The stage was therefore set at the turn of the century for the use of arterial substitutes to bridge gaps when end to end anastomosis was not possible

The first arterial substitutes were solid implants of ivory with a homologous vein transplant placed in the tube as a lining (Nitze, 1897) a method the principle of which was revived in 1942 when Blakemore, Lord and Stefko introduced their *method of inserting vein transplants inside vitellium tubes* Even earlier Abbe (1894) had used a simple glass tube and later a variety of other solid tubes were used but as might be ex pected, the failure rate was very high and today the only use of a

solid and rigid implant is to carry a valve as recommended by Hufnagel, Rabil, Harvey and McDermott (1954) The accompanying Table 1 lists some of the materials which have been

TABLE 1 Materials used as arterial substitutes

Type	Group	Material
1 AUTOGENOUS GRAFTS		Arteries (Hopfner 1903)
		Veins (Lexer 1907)
		Skin (Javid *et al* 1955 Horton *et al* 1956)
		Pericardium (Sako, 1951)
2 HOMOLOGOUS TRANSPLANTS	Arteries	Fresh (Hopfner 1903)
		Refrigerated (Carrel 1901)
		Frozen (Hufnagal and Eastcott 1950)
		Freeze dried (Hyatt *et al* 1951)
	Veins (Watts 1907)	
3 HETEROLOGOUS TRANSPLANTS	Arteries (Carrel 1907)	
4 IMPLANTS	Solid	With vessel inlay (Nitz 1897 Blakemore Lord and Stefko 1942)
		Without vessel inlay
		Glass (Abbe 1894 Tuffier 1915)
		Aluminium and gold (Carrel 1912)
		Silver (Tuffier 1917)
		Leucite (Hufnagel 1947)
		Polyethylene (Donovan 1949)
	Pliable and Pervious	Vinyon N (Voorhees Jaretzki and Blakemore 1952)
		Orlon (Hufnagel and Rabil 1955)
		Terylene (Dacron) (Deterling and Bhonslay 1955)
		Multilayered (Shumacker and King 1954)
		Polyvinyl alcohol (Shumway Gliedman and Lewis 1955)
		Crimped nylon (Edwards and Tapp 1955)
		Nylon (Poth *et al* 1955)
		Teflon (Hufnagel 1955)
	Absorbable	Caramel (Carrel 1902)
		Fibrin (Swenson and Gross 1947)

used as arterial substitutes with the name of whom I believe was the earliest user This list is not complete but it gives a general picture of the variety of arterial substitutes

The most usual arterial substitutes in clinical practice are autogenous venous grafts and homologous arterial transplants

Because of the obvious limitations of supply, autogenous arterial grafts, the ideal material, cannot be used to replace medium sized or large arteries in man, although Sandblom, Moren, Norden, Idbohm and Sandegard (1953) have devised a method of making wide autogenous arterial grafts from narrow vessels An autogenous venous graft has many theoretical advantages over a homologous arterial transplant

Autogenous Vein Grafts

These being autogenous survive transplantation and form a permanent living union with the tissues of the host but, as Table 2 shows, our results with vein grafts in patients have been unsatisfactory However, this may well be due to the type

TABLE 2 Follow up of 228 arterial reconstruction operations (minimum follow up 1½ years)

Operation	Number of patients	Thrombosed early and late	Dead	Patent today
Direct suture	25	1	3	21
Thrombo-endarterectomy	24	5	0	19
Autogenous vein	25	15	4	6
Homologous artery	109	34	10	65
Plastic cloth	19	0	5	14
Polyvinyl alcohol	26	9	5	12

of patient for which we have used them and it is possible that if we had used them purely for traumatic aneurysms as Weglowski did our results might be as good as his, when in 1925 he recorded his follow up on fifty one such grafts inserted during the first war and forty were clinically patent Pringle in 1913 in Britain recorded two successful vein grafts and, as stated in Table 1, they were first used in patients by Lexer (1907)

Histological examination shows that the wall of a vein when subjected to a flow of blood at arterial pressure thickens due to an increase in the fibrous tissue This only happens if the vein is supported by the tissues of the host If the vein is inserted into a large artery within a body cavity it is unable to withstand the increased pressure and it ruptures in a high proportion of experimental animals

Homologous Arterial Transplants

These are today the most widely used arterial substitutes The first essential is to have an arterial bank and, as given in Table 1 there are three main methods of storing arteries Refrigeration has the disadvantage that it is unsafe to use arteries which have been stored in this way for more than six weeks The other two methods work well in clinical practice and we use both the frozen and freeze dried product today at St Mary's Hospital

There was at one time considerable controversy as to whether an artery should be stored in a viable state or not but it is now known that homologous transplants do not survive in the host and that an arterial transplant merely acts as an inert tube into which the tissues of the host grow and so survival during storage which was thought to be important by Gross Bill and Peirce (1949) is now taken to be unnecessary This has the additional advantage that the arteries can be sterilized by any method which does not denature the proteins The following methods have been used to sterilize arteries before storage formalin (Levin and Larkin, 1908) irradiation (Meeker and Gross 1951), ethylene oxide (Hufnagel Rabil and Reid 1953) beta propriolactone (Szilagyi *et al* 1954) and antibiotics plus freeze drying (Fisher Adams Wilde and Fisher 1956) The first of these methods with formalin is not very satisfactory because the proteins are denatured to an appreciable extent but the others have worked well

The fate of a homologous arterial transplant is of some interest and since 1954 we have altered our views on this subject as human material has become available for study (Rob 1954) At one time it was thought that the tissues of the host grew in to replace the donor artery intima grew out from the ends of the host vessel to provide a new lining for the transplant and fibrous tissue grew in to replace the media and adventitia eventually all that remained were the elastic fibres and these extracellular structures persisted for years in an apparently unchanged state We now think that this does not happen in man We have had the opportunity of studying a number of human arterial transplants which have been in position for several months or

even more than a year since insertion, all were much longer than the usual experimental arterial transplants which have formed the basis of former studies These human transplants show only a little replacement by the tissues of the host After one year the intima had grown out 1 to 2 cm from the end of the host artery, in the gap between this true intima the trans plant is covered by fibrin and the media and adventitia have died, but most of these layers have persisted as dead donor tissue With the passage of time there is an increasing degree of replacement with fibrous tissue, particularly near the anastomosis and on the adventitial surface of the transplant Lastly, the elastic laminae remain This means that a human arterial transplant more than 4 cm long consists over its central portion of a layer of fibrin, two layers of dead and almost amorphous donor artery, the elastic laminae and some capillary tufts and areas of fibrous tissue where the host tissues have grown in *The preservation of the elastic fibres is the main reason why* homo logous arterial transplants have proved successful in clinical practice and why the transplants do not become aneurysmal, for as Creech, De Bakey, Self and Halpert (1954) have shown with hererologous transplants the elastic fibres fragment rapidly and aneurysms frequently form

Plastic Implants

As Table 1 shows, a variety of implants have been used to bridge defects in arteries but the only ones which have proved to be of real value are the pliable and pervious plastic materials such as the plastic cloths and sponges These have given en couraging results since they were first used by Voorhees Jaret zki and Blakemore in 1952 Owen (1956) has listed the require ments of the ideal plastic implant for arterial surgery as follows

1 High tensile strength in wet state
2 Durability
   (a) Resistance to chemical action, e g hydrolysis
   (b) Resistance to mechanical action e g work hardening
3 Smooth inner lining (non thrombogenic)
4 Flexibility

5 Stretch
6 Not kinkable
7 Ease of suturing
8 Ease of manufacture
9 Porous wettable outer surface allowing fibroblast permeation yet leakproof
10 Non carcinogenic
11 Not producing allergy or sensitivity reactions
12 Ease of sterilization
13 Freedom from impurities

In clinical practice today a number of materials are in use which fulfil many of these criteria but none is ideal Polyvinyl alcohol sponge prepared in the manner recommended by Shumway, Gliedman and Lewis (1955) scores more points than any of the other materials when compared with the ideal requirements listed but unfortunately it has not stood up to the most important test of all the test of time If it is not made sufficiently thick (at least 4 mm) it may rupture and its very elasticity may be a fault leading to delayed rupture My opinion is that the ideal plastic for arterial replacement has yet to be introduced into clinical practice that polyvinyl alcohol sponge was promising but that woven or knitted plastic tubes of vinyon N orlon terylene or nylon yarn are more satisfactory These plastic cloths have been in use for seven years in experimental animals the sponge tube for only three years As Table 2 shows, the thrombosis rate with the latter has been very high

A large variety of materials work well in the aorta and the decision as to which is the best will have to wait until at least another ten years have elapsed because delayed ruptures may render unsuitable an initially promising material Another problem is that some of these materials may prove to be carcinogenic Oppenheimer Oppenheimer Danishefsky Stout and Eirich (1955) have found that malignant tumours were induced in rodents by *subcutaneously embedding films of the following* polymers cellophane dacron polyethylene polyvinyl chloride silastic pliofilm nylon polymethyl methacrylate polystyrene,

saran, ivalon, kel f, teflon and silk This complication has yet to be reported in man, nevertheless this work raises the whole question of the desirability of using plastics, not only as arterial substitutes but for any surgical procedure where they are embedded in the body, at least until this question is settled

The behaviour of plastics as arterial substitutes in the medium sized vessels is of particular interest In our experience nearly all plastic implants of the following materials—polyvinyl alchohol sponge, orlon cloth and crimped nylon—thrombose within one year when inserted into vessels distal to the bifurcation of the common femoral artery of man On the other hand, we feel that it will not be long before a suitable plastic is produced which will function for prolonged periods in vessels of this size In this connection other authors have reported a number of successes with various plastic implants into the femoral or popliteal arteries, but we have not been so fortunate The reasons why these small implants thrombose whilst the same material gives a good result in the aorta are, we think, multiple, but the following are in our view of special importance firstly the size of the vessel, secondly, the rate of blood flow, thirdly, the relative immobility of the aorta compared with a limb artery particularly near to a joint fourthly, the tendency to kinking in a limb vessel after an implant has been inserted (this usually occurs in the host artery a few inches away from the suture line) and lastly the increased risk of wound infection after a plastic implant has been inserted, particularly to a region such as the groin However, it is of special interest that in our experiments polyvinyl alcohol sponge tubes have remained patent for more than two years in the femoral arteries of four out of eight dogs a result we have been unable to reproduce in man

The microscopical and macroscopical findings of a plastic arterial implant after it has been in position for some weeks, months or years will now be described A few hours after insertion the surface of the implant has become covered with a layer of fibrin or collagen like material containing multiple layers of flattened cells, the thickness of this lining layer varies from 1 millimetre or less to 2 or 3 millimetres and appears to

be thicker when the tube has been made of polyvinyl alcohol sponge rather than one of the plastic cloths To the naked eye this inner lining resembles intima but it is not a true endothelium, from each end of the host artery true endothelium grows out but it does not do more than cover the immediate region of the suture line One possible disadvantage is that this inner layer is not necessarily firmly adherent to the plastic—it could become detached, this has been cited as a cause of late thrombosis Outside the plastic a thick layer of fibrous tissue is laid down, in the case of the plastic cloths this outer layer grows into the implant and it is firmly incorporated into the host but with some other plastics the ingrowth is not so satisfactory and the implant lies almost free within a fibrous sheath

It is of interest that the fibrous or scar tissue which encloses the implant shows no constricting tendency In our experience postoperative constriction in these patients is either due to faulty technique so that narrowing occurs at the anastomosis or to the formation of fresh plaques of atheroma on the host vessels close to the suture lines The most likely reason why this scar tissue does not constrict the implant is that the blood at arterial pressure is sufficient to prevent it

The permeability of these plastic implants is of considerable importance If they are completely impermeable there is of course no leakage of blood but the implant is not completely incorporated in the tissues of the host if they are too permeable blood loss during the operation is excessive The best prostheses are sufficiently permeable to allow the escape of a little blood This then clots in the substance of the implant and a scaffolding is formed into which the tissues of the host can grow With some of the plastic cloths the leakage may be excessive this can be reduced by the simple manoeuvre of introducing blood into the prosthesis so that it is preclotted This means that it is well soaked in fibrin and blood clot before it is used and many of the holes will have become plugged

The length of a plastic prosthesis or homologous arterial transplant which will work satisfactorily has been the subject of some research and it appears that long transplants twenty five or more centimetres work just as well as short transplants

In other words, the diameter is of more importance than the length This is an interesting finding and the explanation may well lie in the way that the lining membrane forms McCune and Blades (1951) conclude that the success of aortic transplants is in no way related to their length, except that long grafts on the whole fare better than short ones In the case of a homologous arterial transplant the endothelium disappears within a few days Klotz, Permar and Guthrie (1903) thought that there was rapid regeneration of endothelium by inward growth from the ends of the transplant, and McCune and Blades thought that the new endothelium came from the recipient vessel at the ends of the transplant and possibly from metaplasia of fibroblasts in the transplant itself The fact that long grafts work better if anything than short ones lends support to the opposite view that the lining is provided by cells of the blood stream which are deposited on to its surface, in fact that the initial covering is really altered and modified blood clot

The Long term Behaviour of Arterial Substitutes

This is the great unsolved question In the case of autogenous vein grafts it is known that they function satisfactorily as arterial substitutes for twenty or more years once they have passed through the difficult initial period, when most of the thromboses occur, but no such knowledge is as yet available for the other forms of arterial substitutes homologous arterial transplants have only been used in man since 1949 and the longest follow up is therefore only seven or eight years, in the case of plastic implants the solid tubes which were the first to be used were unsatisfactory and so the longest follow up on a material of reasonable value is only since 1952 or four years In spite of this short period of study certain conclusions can be reached at least of a temporary and provisional nature These are that in the present state of our knowledge autogenous vein grafts and homologous arterial transplants are the safest for use in the peripheral arteries of patients that of the plastic implants the various cloths, vinyon nylon orlon and terylene are satisfactory for use in the human aorta that polyvinyl alcohol sponge gave good

initial results but the longer term behaviour was most unsatis factory

Coleman, Deterling and Parshley (1955) have made a study of homologous arterial transplants eighteen months to four and a half years after insertion In all the functional result was excellent and no aneurysms were noted However, calcification and marked thinning were observed locally Microscopical examination revealed condensation of the graft with complete loss of cellular elements and a thinning of the internal and external layers which had developed from the fibroblasts of the recipient In particular there was marked degeneration of the elastic fibres of the transplant These authors also found that calcium salts had been deposited in the graft, a finding which has been noted by several other workers These findings show that this subject cannot be viewed with complacency until really long term studies are available, but in the interval we should continue to use these materials because of the lives they save and the symptoms they relieve In the case of children the special factor of the growth of the transplant must be considered, particularly when it has been used as part of an operation for correction of a coarctation of the aorta in a young child The evidence is that a homologous arterial transplant will increase in size when implanted into the aorta of a growing animal but that the rate of growth is less than that of the host vessel so that a narrowed segment results If continuous as opposed to inter rupted silk sutures are used a stricture forms at the site of the anastomosis

Conclusion

A large variety of materials have been used as arterial sub stitutes The most satisfactory are autogenous veins and homo logous arteries Of the various plastic implants good results have followed the use of woven plastic cloths There is evidence that some of these plastic materials may be carcinogenic when implanted into the mammalian body

Apart from autogenous vein grafts little is known about the long term behaviour of these arterial substitutes Homologous arterial transplants were first used in patients in 1949, the plastic cloths in 1952 and polyvinyl alcohol sponge in 1954

At least another fifteen years must elapse before the long term behaviour of even homologous arterial transplants in the human body can be satisfactorily assessed, but in the interval the clinical use of arterial substitutes is justified because of the lives they save and the symptoms they relieve

REFERENCES

ABBE R (1894) *New York Med J* **59** 33

BLAKEMORE A H, LORD J and STEFKO P (1942) *Surgery* **12** 488

CARREL A (1902) *Lyon Med* **93** 859

CARREL A (1907) *Bull Johns Hop Hosp* **18** 18

CARREL A (1907) *J exp Med* **9** 226

CARREL A (1910) *J exp Med* **12** 460

CARREL A (1912) *Surg Gynec Obstet* **15** 245

COLEMAN C C DETERLING R A and PARSHLEY M S (1955) *Surgery* **27** 64

CREECH O DE BAKEY M E SELF M and HALPERT B (1954) *Surgery* **36** 431

DETERLING R A and BHONSLAY S B (1955) *Surgery* **38** 71

DONOVAN T J (1949) *Ann Surg* **130** 1024

EDWARDS S T and TAPP J S (1955) *Surgery* **38** 61

FISHER B ADAMS C WILDER R and FISHER E R (1956) *Ann Surg* **143** 73

GROSS R E BILL A H and PEIRCE E C (1949) *Surg Gynec Obstet* **88** 689

GUTHRIE C C (1912) *Blood Vessel Surgery and Its Applications* Arnold London

HOPFNER E (1903) *Arch J klin Chir* **70** 417

HORTON C CAMPBELL F CONNAR R SMITH A and PICKRELL K (1956) *Surgery* **39** 926

HUFNAGEL C A (1947) *Arch Surg* **54** 382

HUFNAGEL C A and EASTCOTT H H G (1950) *Amer Coll Surg Surg Forum* 269

HUFNAGEL C A RABIL P J and REID L (1953) *Amer Coll Surg Surg Forum* 162

HUFNAGEL C A RABIL P J HARVEY W P and MCDERMOTT T F (1954) *Surgery* **35** 673

HUFNAGEL C A and RABIL P J (1955) *Arch Surg* **70** 105

HUFNAGEL C A (1955) *Surgery* **37** 165

HYATT G W TURNER T C and BASSETT C A L (1951) *Navy Med News Ltr* **18** 11

JABOULAY M and BRIAN E (1896) *Lyon Méd* **81** 97
JASSINOWSKY A (1889) Quoted by Horsley J S in *Surgery of Blood Vessels* Morsby St Louis 1915
JAVID H DYE W S GROVE W J and JULIAN G C (1955) *Ann Surg* **142** 613
KLOTZ O PERMAR H H and GUTHRIE C C (1923) *Ann Surg* **78** 305
LAMBERT M (1764) *Med Obs and Inq Soc Phy Lond* **2** 360
LEVIN I and LARKIN J H (1908) *Proc Soc exp Biol Med* **5** 109
LEXER E (1907) *Arch J klin Chir* **83** 459
McCUNE W S and BLADES B (1951) *Ann Surg* **134** 769
MEEKER I A and GROSS R A (1951) *Science* **114** 283
MURPHY J B (1897) *Med Record* **51** 73
NITZE M (1897) *Zentral J Chir* **24** 1042
OPPENHEIMER B S OPPENHEIMER E T DANISHEFSKY I STOUT A P and EIRICH F R (1955) *Cancer Research* **15** 383
OWEN K (1956) *Proc Roy Soc Med* **49** 340
POTH E J JOHNSON J K CHILDERS J H and GUY R S (1955) *Amer J Surg* **89** 1196
PRINGLE H (1913) *Lancet* **i** 1795
ROB C G (1954) *Lancet* **ii** 255
SAKO Y (1951) *Surgery* **30** 148
SANDBLOM P MOREN A NORDEN G IDBOHM E SANDEGARD E and DAHLBACK O (1953) *Acta chir scand* **106** 209
SHUMACKER H B and KING H (1954) *Surg Gynec Obstet* **99** -87
SHUMWAY N E GLIEDMAN M L and LEWIS E J (1955) *Surg Gynec Obstet* **100** 703
SWENSON O and GROSS R E (1947) *Surgery* **22** 137
SZILAGYI D E OVERHULSE P R and LO GRIPPO G A (1954) *Clin Res Proc* **2** 108
TUFFIER M (1915) *Bull Acad de Med Paris* **74** 455
TUFFIER M (1917) *Bull Soc Chir* **43** 739
VIDAL M E (1903) *Press Med* **11** 747
VOORHEES A B JARETZKI A and BLAKEMORE A W (1952) *Ann Surg* **135** 332
WATTS S H (1907) *Bull Johns Hop Hosp* **18** 179
WEGLOWSKI R (1925) *Zbl Chir* **52** 2241

V

Lung Function Tests an Assessment of their Usefulness

J C GILSON

Historical

TESTS of renal function have been in general clinical use for more than thirty years Why is it only in the last five years that clinical tests of lung function have rivalled those of the kidney for complexity and comprehensiveness?

Professor G C Douglas (1954), in a review of pulmonary function in man, has shown how physiologists in the last century and the early part of this were mostly interested in disentangling the factors controlling ventilation They studied the effects of oxygen and carbon dioxide and later pH, following the discovery of the production of lactic acid during muscular contraction by Fletcher and Hopkins (1907) Much of the interest was in factors controlling ventilation rather than in the function of the lungs themselves

There have, of course, been exceptions, starting with the first measurements of lung volume by Humphrey Davy (1839) There is the classical work on vital capacity by Hutchinson and at the end of the first quarter of this century there were many clinical studies of the vital capacity by West and Myers and others The little help that this single test brought may have *prejudiced* the clinician against all tests of lung function, but if the number of papers now appearing annually on this subject is a guide, the prejudice i a relic of the past

The Second World War brought great impetus to the study of lung function, particularly of normal subjects exposed to

extremes of ambient pressure—high in the Navy and low in the Air Force It also brought new techniques Methods of rapid physical gas analysis and means of measuring small and rapidly fluctuating pressure at a distance with the new fine plastic tubes

Since the war an increasing proportion of the literature on lung function comes from the clinical departments The abnormal has been used to clarify the function of the normal Five years ago in this series Dr K W Donald (1953) gave a full account of current methods of assessment of respiratory function but it is an index of the rapidity of advance in this subject that the details of the mechanics of breathing were not men tioned Comprehensive reviews of the whole subject have appeared since (Gaensler 1955) I shall therefore, select some of the most recent advances and also attempt to show how studies of lung function are contributing to various aspects of medicine

DEFINITIONS

Table 1 shows a classification of lung function By common usage tests of lung function apply principally to tests concerned with the first two—*ventilation* and *gas exchange* This is, of course an arbitrary separation The pulmonary circulation is just as much a part of its function but I shall be concerned with the first two only

TABLE 1 Lung function

VENTILATION
Size
Gas distribution
Mechanics
GAS EXCHANGE
Ventilation/perfusion ratio
Diffusion
CIRCULATION
OTHER FUNCTIONS
Metabolic etc

There are other aspects of lung function at present inacces sible for study in man but not therefore unimportant We know very little of the normal function of its nervous system and virtually nothing of its derangements There is the metabolism

of the lung itself, a study of which might increase our understanding of the degenerative processes of emphysema and age There are other remarkable functions, for example, the ability of the lungs to separate off the leukocytes and then release them again—a function at present without obvious usefulness (Bierman *et al*, 1952)

Opinion on the usefulness of tests of lung function must inevitably be a personal one—for example, tests to determine the part the red cell membrane contributes to the total diffusion barrier between the haemoglobin and the alveolar gas are unlikely to be of use to the thoracic surgeon In contrast, overall tests of ventilatory capacity may be of considerable use to the surgeon but of little use to the physiologist However, people with such different interests are now using tests of lung function that it is interesting to consider the diversity of their usefulness

Recent Advances

In the last five years there has been a rapid advance in methods for separating and measuring the components controlling ventilation and gas exchange One result of this is that a fully equipped laboratory can now provide the clinician with as many as one hundred test results on an individual As a further result there is a growing need for a more comprehensive and quantitative classification of the syndromes of derangement of function revealed by these tests

Lung Volume

Before the mechanics of breathing are discussed mention must be made of a new approach to the measurement of lung volume Although it is now recognized that the residual volume, as a proportion of the total lung capacity can be considerably raised in conditions other than emphysema and is, therefore, a less specific test than was at one time thought its measurement and that of the functional residual capacity is still important The gas replacement and dilution methods in general use only measure the gas in direct communication with the trachea Bedell, Marshall, DuBois and Comroe (1956) have recently developed a novel method of measuring the total gas volume

within the thorax It is a pneumometric method which measures not only the gas within the lung which is in communication with the trachea, but also cysts and volumes extremely poorly ventilated as may occur in some cases of emphysema The difference between the volumes measured by the gas dilution and pneumometric method, therefore, gives a new measurement—the volume of the lung which is very poorly ventilated An advantage of the new method is that it is much quicker, a disadvantage is that more complex equipment is required Table 2 shows that the volumes by the two methods

TABLE 2 Comparison of thoracic gas volume and functional residual capacity

Subjects	Normal	Emphysema
No	9	11
Age	20–43	47–81
F R C (7 min N_2 washout)	2 97 l	4 53 l
Thoracic gas volume (Pneumometric body box)	2 97 l	5 62 l
Difference	—	1 09 l

(Ref DuBois *et al* 1956 Bedell *et al* 1956)

are the same in young normal subjects but are different in cases of advanced emphysema It will be interesting to see comparisons with those of comparable age and with the results of the alternative helium dilution method of measuring the functional residual capacity

Mechanics of Ventilation

Specific Tests

Although it is forty years since Rohrer (1915) made the first theoretical analysis of the forces and resistances concerned in breathing it is only recently that measurement of the component resistances has become possible Figure 1 shows diagrammatically the forces and resistances opposing inspiration and the components which are now measurable

The elastic restorative force is dependent on the degree of inflation (i e the volume) and this force increases gradually with the depth of inspiration the energy of stretching the spring

being stored in the lung and thoracic wall It is now usual to express the relationship of volume and this elastic force as compliance—l /cm H_2O This measurement of compliance can only be made when no air is flowing

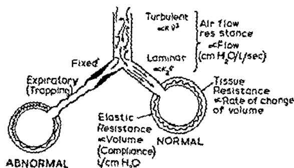


FIG 1 Mechanics of breathing

The resistance caused by the air flow and the viscous resistance within the tissue are also shown diagrammatically The air flow within the lung is part laminar, part turbulent, therefore, the pressure drop between the mouth and the alveoli is proportional to something less than the square of the velocity (approx $V^{1\ 66}$) To measure the combined resistances of the tissues and the air flow it is necessary to record the pressure outside the alveoli but within the chest wall simultaneously with the air flow at the mouth It has now been established that within certain limits (of depth and rate of breathing) the pressure changes in the oesophagus run parallel with those in the intra pleural space, thus by plotting the oesophageal pressure against the flow at the mouth or the change in volume it is possible to calculate air flow resistance and compliance Separation of that part of the resistance caused by the tissues from that caused by the air flow requires a measurement of the pressure drop between the inside of the alveoli and the mouth DuBois, Botelho and Comroe (1956) believe that they have succeeded in doing this using a modification of the pneumometric method already mentioned The resistances of the two components have been studied in a variety of conditions and it

appears probable that the air flow resistance usually forms the major part (Marshall and DuBois, 1956)

The compliance has been found to be reduced, that is to say the lungs are more rigid than normal in poliomyelitis, in fibrosis of the lungs and thorax, in congestive heart failure and also under general anaesthesia A small increase of compliance has been noted in emphysema but the effect on the average is not large (Attinger, Goldstein and Segal, 1956)

In normal subjects the air flow resistance is similar during inspiration and expiration during normal breathing In chronic obstructive emphysema there is usually a gross difference in the resistance in the two phases of respiration with a marked increase on expiration, particularly at high flows, this is believed to be due to collapse of the bronchioles causing trapping of the air (Mead and Whittenberger 1953, Fry, Ebert, Stead and Brown, 1954) In asthma the resistance is raised during both phases Aerosols of histamine and acetylcholine can be shown to produce two- or threefold increase in air flow resistance Anti spasmodics particularly in cases of asthma, have the reverse effect

Measurements of the resistances have enabled calculations of the energy dissipated in breathing to be made, this energy is greatly increased in emphysema and in asthma It has been shown that when conditions within the lung change, for example when it becomes more rigid as in pulmonary con gestion the rate and depth of breathing are adjusted so as to keep the work of breathing at a minimum (Marshall McIlroy and Christie, 1954)

Overall Tests[1]

These tests of the separate components of the mechanics of breathing have not replaced the overall tests of the ventilatory capacity such as the maximum voluntary ventilation and the single forced expiration test—the results of which are affected by changes of air flow and tissue resistances, compliance of chest and lung muscular force and the size of the lung The

[1] This lecture was given before the new terminology for measurements of ventilatory capacity (Gandevia and Hugh Jones 1957) was agreed

great simplicity of these tests has led to their being used more widely than any other tests of lung function at the present time

Both tests have excellent discrimination—that is, the variability of results in a given normal individual is small (S E of mean of three readings approximately 3 l /min) compared to the range between normal and abnormal, they are, therefore, sensitive tests The tests give similar but not identical information

The maximum voluntary ventilation—the volume of air breathed (expressed in l /min) over a period of 15–20 seconds of maximum voluntary hyperventilation—increases with rate of breathing up to at least 80 per minute (McKerrow, 1955) Some observers therefore prefer to select a particular rate of breathing for example the upper limit of that normally observed on exertion—about 40 per minute To avoid erroneous results the apparatus used must have a very low flow resistance and inertia (Bernstein, D Silva and Mendel, 1952) It is a test which is influenced by both inspiratory and expiratory resistance within the lung and is, therefore, a sensitive measure when both are raised Nevertheless, on account of the instrumental difficulties and the discomfort caused, there has been increasing trend to use the single breath method as an alternative test

There have been two approaches to this single breath analysis Some have sought to measure the maximum voluntary ventilation indirectly by its use Instead of adding up the total volume breathed in and out in the 15 or 20 seconds, the volume during a single expirate over the first 0 75 second is recorded The assumption is then made that the inspiratory and the expiratory time periods are the same and the minute volume is calculated assuming a breathing rate of 40 per minute

When this is done the correlation between the direct and indirectly estimated maximum voluntary ventilation is very high (Kennedy, 1953) Even when the rate of breathing during the maximum voluntary ventilation is not controlled at the rate of 40 per minute the correlation is still about 0 8 Figure 2 shows the results of using both tests on a whole colliery population of 250 men The pattern of the relation of X ray category of simple pneumoconiosis and age revealed by the two tests is essentially the same

An alternative approach has been to obtain as much information as possible from the shape of the spirogram during forced expiration In emphysema and asthma the initial flow rate is low and the curve long drawn out In cases of congestion,

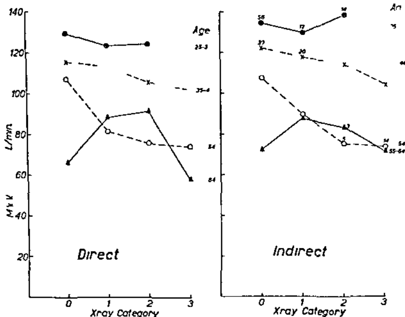


FIG 2 The relation of age and radiological category of simple pneumoconiosis to ventilatory capacity measured in two ways in a group of working coal miners (ILO Classification 1950)

The direct maximum voluntary ventilation was measured over a period of 15 seconds maximum voluntary hyperventilation at a breathing rate chosen by the subject

The indirect maximum voluntary ventilation was estimated by multiplying the volume expelled in the first 0 75 second of a forced expiration by 40

particularly in mitral stenosis the volume is small and the end of the curve sharply flexed These are qualitative observations and there is as yet no standardization of the quantitative measurements from these curves The volume expelled over periods of 0 5, 0 75 1 2, 3 seconds is used If a very short period is used the result is liable to error from inaccuracy of

timing If long periods are chosen it becomes insensitive at differentiating between subjects who can expel their whole vital capacity within about a second Between 0 75 and 1 second is probably the best compromise Figure 3 shows the high correlation

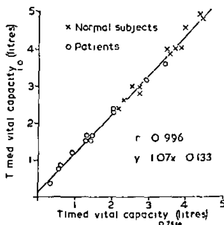


FIG 3 The relation between the volume of gas expelled during the first 0 75 and 1 second of a forced expiration following a maximum inspiration The patients had pneumoconiosis emphysema and bronchitis

between the volume expelled in 0 75 and 1 second over a wide range of absolute volumes Recently Leuallen and Fowler (1955) have suggested that the flow rate over the middle half of the whole volume is the index correlating most closely with other physiological assessments of emphysema

Some, following Gaensler s and Tiffeneau s suggestions, have used the percentage of the vital capacity expelled in one second as an index of the severity of obstructive pulmonary disease In normal subjects between 70 and 90 per cent is expelled within one second In obstructive conditions the percentage may be reduced to 50 or less

In general the measurement of the volume during the forced expiration is not readily falsified by the type of apparatus used and it has been shown that the volume expelled in periods of 0 5 second and over is the same whether recorded on a spirometer or integrated from the record of an inertia free pneumotachygraph When more is known about factors affecting this

test and its opposite the flow during inspiration they are likely to provide even more useful information than at present, but there is need for standardization of nomenclature describing them (Gandevia and Hugh Jones, 1957)

DIFFUSION

Unevenness of lung ventilation is relatively easy to detect and measure Inequalities in blood flow in different parts of the lung have been less easy to detect or measure When the ventilation and blood flow alter together so that their ratio remains the same, no important functional derangement occurs When, however, there is a change of ratio, as for example a decrease of ventilation relative to the blood flow, the blood leaving that part of the lungs is deficient in oxygen and there is no compensation in the other areas where the ventilation exceeds the blood flow because the shape of the oxygen dissociation curve results in no appreciable increase of saturation in the hyperventilated area The measurement of this ventilation perfusion ratio averaged for the whole lung is possible but is technically difficult and no great simplification of the ingenious methods described by Riley and Cournand (1949) has yet been published However Hugh Jones and his colleagues at the Postgraduate Medical School have recently demonstrated the great possibilities opened up for its measurement by using the mass spectrometer recording three gases simultaneously [1]

Even if there is a normal ventilation perfusion ratio the blood leaving the lung may still be undersaturated if the barrier to diffusion of gas from the alveoli to the haemoglobin is increased There is no direct method of measuring the thickness of the alveolar wall or the surface area of the lung in life but it is possible to measure the diffusing capacity of the lung as a whole

$$\text{Diffusing capacity for } O_2 \text{ or } CO = \frac{\text{cc uptake of gas/min}}{\text{Mean alveolar partial pressure} - \text{Mean capillary partial pressure}}$$

The avidity with which carbon monoxide and oxygen are taken up by the haemoglobin results in the alveolar membrane

[1] See West Fowler Hugh Jones and O Donnell 1957

being the principal factor limiting (at least up to levels of moderate exertion) the rate of uptake of these two gases

Carbon monoxide, first used for this purpose by Krogh in 1915, has the advantage over oxygen that its affinity for haemoglobin is so great that within the few minutes necessary to carry out the measurements no important degree of back pressure of gas arises along the whole length of the lung capillaries thus in the equation the capillary partial pressure becomes zero and we are left with the need to measure only the uptake of gas per minute and the mean alveolar carbon monoxide partial pressure The mean alveolar pressure cannot be measured directly

TABLE 3 Diffusing capacity of the lungs for carbon monoxide cc /min /mm Hg

	Method	No	Mean age	Resting	Exercise
Krogh M (1915)	Breath holding	15	34	30 (19–41)	—
Bates *et al* (1956)*	Breath holding	6	30	17 (12–22)	—
Filley *et al* (1954)†	Steady state	11	31	17 (13–28)	36 (23–55)
Bates *et al* (1955)‡	Steady state	19	28	25 (11–32)	34 (24–47)

* Alveolar CO estimated from single expirate
† Alveolar CO calculated from arterial CO_2 tension
‡ Alveolar CO estimated from end tidal sample

Indirect methods have been used for obtaining an approximate value and the results of the diffusing capacity for normal subjects agree reasonably well for the recent determinations using different methods (Table 3) In abnormal cases with marked inequality of ventilation the estimate of the mean alveolar carbon monoxide pressure is more in doubt and the exact interpretation of the result less certain, but the test probably provides the best measure we have of the active diffusing surface of the lung

One of the striking findings has been a rise of diffusing capacity with exercise up to about 75 per cent above the resting level About an equal percentage reduction in the diffusing capacity has been noted in conditions with a thickened alveolar membrane or reduced alveolar surface —such as emphysema It is

however possible that the reduction in emphysema is apparent rather than real and accountable for by the marked inequality of ventilation

Uses of Lung Function Tests

To what uses have these tests of lung function been put? As Table 4 shows their use now covers too wide a field to be dealt with adequately in a single lecture so I will select a few examples to demonstrate their uses and limitations

Table 4 Uses for lung function tests

Physiological	Functional dissection of the lung
Medical	Diagnosis
	Management
	Evaluation of therapy
Surgical	Pre-operative assessment
	Assessment of operations
Medico-legal	Assessment of disability
Epidemiology	Aetiology
	Prevention

Physiology

Few would, I think, deny that the greatest use of these tests so far has been to permit a functional dissection of the lung which has greatly increased our understanding of its normal and abnormal action and so provided a scientific basis for the physiological interpretation of diseases of the lung and to a lesser extent the chest wall

Medicine

Diagnosis Medical teaching places great emphasis on diagnosis in pathological terms It is therefore somewhat disconcerting but stimulating to find that results of the tests of lung function rather rarely assist in this aim The reason is simple A number of aetiologically different pathological processes may produce a similar type of functional derangement The tests provide a diagnosis in terms of function but the tests are now so numerous that a description of the physiological state is inevitably lengthy Table 5 shows an abbreviated description of three functional syndromes, but there are several others

In the first syndrome the total lung capacity is reduced the

being the principal factor limiting (at least up to levels of moderate exertion) the rate of uptake of these two gases

Carbon monoxide, first used for this purpose by Krogh in 1915, has the advantage over oxygen that its affinity for haemoglobin is so great that within the few minutes necessary to carry out the measurements no important degree of back pressure of gas arises along the whole length of the lung capillaries, thus in the equation the capillary partial pressure becomes zero and we are left with the need to measure only the uptake of gas per minute and the mean alveolar carbon monoxide partial pressure The mean alveolar pressure cannot be measured directly

TABLE 3 Diffusing capacity of the lungs for carbon monoxide cc /min /mm Hg

	Method	No	Mean age	Resting	Exercise
Krogh M (1915)	Breath holding	15	34	30 (10–41)	—
Bates *et al* (1956)*	Breath holding	6	30	17 (12–22)	—
Filley *et al* (1954)†	Steady state	11	31	17 (13–28)	36 (23–55)
Bates *et al* (1955)‡	Steady state	19	28	25 (11–32)	34 (24–47)

* Alveolar CO estimated from single expirate
† Alveolar CO calculated from arterial CO_2 tension
‡ Alveolar CO estimated from end tidal sample

Indirect methods have been used for obtaining an approximate value and the results of the diffusing capacity for normal subjects agree reasonably well for the recent determinations using different methods (Table 3) In abnormal cases with marked inequality of ventilation the estimate of the mean alveolar carbon monoxide pressure is more in doubt and the exact interpretation of the result less certain, but the test probably provides the best measure we have of the active diffusing surface of the lung

One of the striking findings has been a rise of diffusing capacity with exercise up to about 75 per cent above the resting level About an equal percentage reduction in the diffusing capacity has been noted in conditions with a thickened alveolar membrane or reduced alveolar surface—such as emphysema It is

The second syndrome is that of a normal or increased total lung capacity, a reduced vital capacity with a much raised residual volume and ratio There is a marked impairment of ventilatory capacity, the distribution of the air is markedly uneven, the compliance is normal or increased the air flow resistance raised particularly on expiration, the diffusing capacity is reduced, and the ventilation on exercise is slightly raised The oxygen saturation is reduced The carbon dioxide tension in the blood may be raised This is the functional syndrome of obstructive emphysema In its fully developed form it is a condition in which the clinician, physiologist, the radiologist and pathologist will probably agree in their diagnosis but there is interest and difficulty in its less severe degrees

The last column shows the changes which have been reported as occurring normally with ageing Most of these studies have referred to groups selected as being perhaps fitter than the average for their age Nevertheless the changes with ageing show a remarkable similarity to those in emphysema but are of a lesser degree in all instances We need to know far more about how lung function alters with age in representative samples of the general population We shall then have a clearer idea whether the functional changes of emphysema constitute a separate entity or merely an exaggeration of the normal ageing process

I think we may have in emphysema a position not very dissimilar from that which now exists in the study of blood pressure where the arbitrary subdivision into hypertension and normal tension has given place to an appreciation that there is a continuous increase of blood pressure with age (Hamilton Pickering Fraser Roberts and Sowry 1954) Such a concept of course does not exclude the likelihood that individuals with high blood pressures or the more advanced degrees of functional abnormality of emphysema may have a different prognosis from those at the other end of the distribution Here is a field of usefulness for lung function studies which has only just started and may be expected to illuminate the whole problem of ageing

The effect of ageing on function is perhaps bigger than is often appreciated there is also a rather remarkable linear decline

vital capacity is reduced, the residual volume is normal but percentage is raised, the ventilation on exercise is increased, the ventilatory capacity is normal or slightly reduced, and the distribution within the lung is not grossly uneven, the compliance is reduced but the air flow resistance is normal, the gradient between the alveolar and arterial blood is raised and the

TABLE 5 Syndromes of lung function

	Alv Cap Block	Emphys	Age
SIZE			
Total lung capacity	–	N	N
Vital capacity	–	–	–
Residual volume	N	++	+
Residual volume % T L C	+	++	+
VENTILATION			
Ventilation (Ex)	++	+	N
Distribution	N	– –	–
Ventilatory capacity	N	– –	–
Compliance	– –	+	N
Flow resistance	N	++	+
GAS EXCHANGE			
Alv PO_2 – Art PO	++	+	+
Diffusing capacity (CO)	– –	– –	–
BLOOD			
O_2 saturation (Ex)	– –	–	N
PCO_2	N	+	N

Symbols N = Normal + Raised – Lowered

diffusing capacity is reduced, often markedly, the oxygen saturation falls on exercise This is then the syndrome of a small stiffened lung with relatively even ventilation but a greatly increased diffusion barrier This condition of the alveolar capillary block as it has been called, has been demonstrated in a number of pathological conditions such as beryllium lung, sarcoid, pulmonary granuloma of a variety of types, diffuse infiltration by neo-plastic tissue, and in some unusual cases of pulmonary fibrosis of the Hammen Rich type Physiological findings alone will be of little assistance in deciding which pathological process is responsible

polate linearly the evidence already available, it makes one realize how little functional reserve in kidneys or lung is left by the time we reach the age of a hundred!

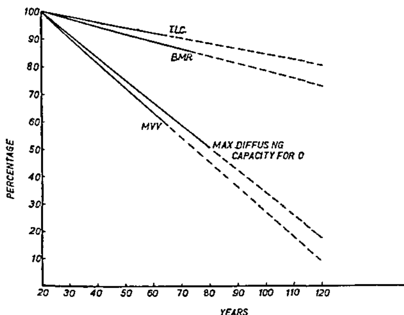


FIG 5 The effect of age and lung function expressed as a percentage of values at age 20
Total lung capacity (T L C.) Gilson *et al* 1955 Needham *et al* 1954
Basal metabolic rate (B M R) Albritton 1954
Maximum voluntary ventilation (M V V) Carpenter *et al* 1956
Maximum diffusing capacity for oxygen Cohn *et al* 1954

These figures about the effect of age have been a digression from the main point—the use of these tests in diagnosis—but have I hope shown how important it is to take age into account before reaching a firm diagnosis Other examples of their uses for this purpose are to establish the existence of a vascular shunt by the absence of a full rise in saturation on giving oxygen but even then the test does not help in deciding whether the shunt is cardiac or pulmonary or again in separating primary from secondary polycythaemia But they can be useful in separating

with age in some cases Figure 4, derived from the findings of Davies and Shock (1950), shows the decline of glomerular filtration rate and effective renal plasma flow with age when expressed as a percentage of the value at age twenty

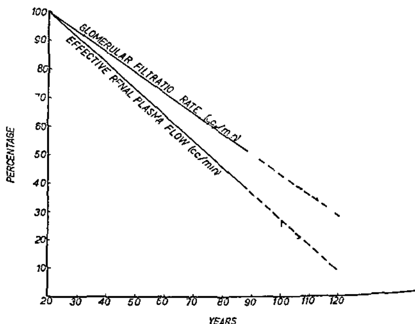


FIG 4 The effect of age on renal function—the change expressed as a percentage of the value at age 20 (Modified from Davies and Shock 1950 70 male subjects aged 24–89)

Figure 5 shows the results of four other tests, three related to lung function The total lung volume and the basal metabolic rate are relatively little affected by age whereas the maximum voluntary ventilation and the maximum diffusing capacity for oxygen fall steeply at about the same rate as the renal tests These are mean values and the coefficients of variations are of the order of 20 per cent Only the maximum voluntary ventilation results refer to a representative sample of the general population Rather similar trends with age have been published in an earlier lecture in this series for the decline of excretion of 17 ketosteroid (Dodds 1953) If one is rash enough to extra

example, in separating the specific from the general effect produced by cortisone and ACTH Objective evidence of a change in diffusing capacity following such therapy has been shown in some cases of beryllium lung, sarcoid and Hammen Rich syndromes, but no changes have been demonstrated in silicosis The maximum voluntary ventilation has been extensively used to compare the many anti spasmodics available for treating bronchospasm It is probable that the air flow resistance is an even more sensitive index for this purpose

Measurement of tidal volume, dead space and alveolar carbon dioxide have been used to study artificial respiration For example Asmussen and Nielsen (1950) showed that in apnoeic subjects Schafer's method only produced a tidal volume of about 180 cc and an alveolar ventilation of less than 50 cc per breath By contrast the Holger Nielsen and the Eve rocking method produced a ventilation three or four times as large Dripps and his colleagues (Price, Conner and Dripps 1954) compared the relative efficiency of pressure and pressure suction respirators in anaesthetized subjects and showed that the ones available in commerce do provide a satisfactory ventilation without serious disturbance of circulation Affeldt and his colleagues (Collier and Affeldt, 1954) have provided accurate comparisons of the efficiency of the abdomino chest and pure chest type curass respirators in cases of poliomyelitis and expressed the results as a percentage of the efficiency of the box respirator

These are only a few of the uses—there are many others such as the assessment of the value of pressure breathing with and without anti spasmodics and to test the effect of breathing exercises In this last example evidence of objective improvement is still unconvincing The changes are principally in the pattern of breathing while conscious control is maintained

SURGERY

If I only just touch on this aspect it must be accepted as evidence of personal bias and ignorance The use of a single test of ventilatory capacity such as the maximum voluntary ventilation must have prevented the production of a number of respiratory

the dyspnoea caused by cardiac failure or severe kypho-scoliosis from that produced by emphysema

Eventually a system of diagnosis based on a classification of the functional abnormality will be possible in chest disease, and it will become really useful when the prognostic significance of the different syndromes has been established

Management

Tests of lung function have a place in the day to-day management of a few chest conditions In patients with ventilatory insufficiency severe enough to cause coma many physicians now regard arterial blood samples as essential in order to establish the oxygen and carbon dioxide tension and pH because the immediate treatment may depend on the findings, for example how to make the best use of oxygen and assisted respiration Similar problems may occur in the management of cases of severe emphysema during post operative recovery from anaesthesia

Another example is poliomyelitis Measurement of tidal volume, the vital capacity and the compliance of the chest wall and lungs can be a help in deciding how much assistance to respiration is needed End tidal samples—estimated for carbon dioxide—have been of use to check the degree of ventilation, for it is now recognized that excess ventilation can be harmful by leading to a respiratory centre adapted to a low carbon dioxide tension so that weening from the respirator becomes difficult

In both these conditions a simple clinical instrument for measuring ventilation would be useful, none is at present available This surprising deficiency will, I hope, be soon remedied by a simple anemometer developed by my colleague, Dr B M Wright (1954) With it the minute volume can be recorded about as easily as the temperature The instrument only weighs a few grams, offers a negligible resistance to breathing, and requires no valves It can also be used for measuring the maximum breathing capacity

Evaluation of therapy

Tests of lung function have been of more use in the objective measurement of the effects of therapy than in management—for

usually two quite separate issues (1) Is the claimant disabled? (2) Is the disability—if present—the result of, for example, pneumoconiosis?

The second question is rarely convincingly answered by examining the individual because the changes of function are non specific in most types of pneumoconiosis It is only by examining representative samples of all those with a particular X ray category of pneumoconiosis (and allowing for the effects of age) that the average disability actually caused by the pneumo coniosis can be accurately determined This essential infor mation is still lacking for many types of pneumoconiosis and it is too often assumed that because there are X ray changes the disordered function must be the result of these changes Only in advanced cases may it be easy to decide but then the tests are not really needed

The first question is also often not easy to answer by function tests because the disability is usually one of excessive dyspnoea on exertion which is a subjective sensation and one which tests of lung function have not been able to measure objectively and simply in individuals with great accuracy As Dr J H Comroe so aptly pointed out recently our understanding of the mech anism of dyspnoea is still so incomplete that we prefer to measure something else which can be done more easily[1]

There is a wide variation in normal values of most lung func tion tests, even after allowing for age, so that the interpretation of small or even moderate deviations from the average is un certain It is a real difficulty here that the law does not recognize the existence of the normal scatter about the average values The expression of the results in terms of percentage of pre dicted normal rarely adds much to the information because predictions are associated with a large possible error—often conveniently overlooked[1] The tests may however sometimes establish that function is normal in those claiming compen sation

While the use and interpretation of tests of lung function for medico-legal purposes are more difficult than appears at first sight, the alternative of relying principally on the comments of the claimant may also be unsatisfactory, particularly when

cripples and enabled numbers of others to benefit from operations—the advisability of which was in some doubt on account of the difficulty of assessing respiratory reserve clinically

The assessment of each lung separately by broncho spirometry, although first used in 1932 by Jacobaeus, has not come into routine use, although it has been of value in special cases—for example, cases requiring decortication Even the latest catheters have a resistance to air flow about three times the bronchial air ways so that measurements of maximum ventilatory or diffusing capacity are not possible

Much of the information obtained from broncho-spirometry about the relative function of the two lungs can also be obtained by sampling the gas in the trachea and bronchi simultaneously through fine tubes as suggested by Armitage and Brian Taylor (1956) The use of the mass spectrometer to analyse simultaneously and continuously the gases from these fine catheters may bring partition spirometry into routine use

As in the medical field, more use has been made in surgery of tests of function to evaluate the effect of treatment Different operative procedures can be compared objectively I will take one example About ten years ago the high residual volume in the remaining lung after pneumonectomy led some surgeons to advise thoracoplasty to prevent the distension of the remaining lung—a distension assumed to be similar functionally to the changes in emphysema However, pre- and post operative studies of thoracoplasty showed how much it often reduced the ventilatory capacity especially when scoliosis developed, and this reduced still further the respiratory reserve These studies of function have influenced opinion towards more conservative measures The small groups with pre and post operative functional studies are gradually being enlarged in numbers and duration now that function studies are used routinely in some centres (Hirdes and Bosch 1955)

Medico Legal

Perhaps more tests of lung function have been done to less purpose in this field than any other This may seem strange It might be thought that tests would be most useful here There are

compensated), make one wonder whether their dyspnoea esti mated clinically was not conditioned more by compensation than by lung dysfunction

Epidemiology

Lastly we come to uses for these tests in epidemiology This section should, of course, really come first for it concerns aetiology and prevention, but it happens to be the field in which they have been least used so far

In epidemiology of chest disease we are concerned with the effect of heredity and environment on function We are thus concerned with the results for groups of people and individual variations are less important, but other difficulties arise If the refusal rate in the groups studied is to be kept acceptably low the tests must be short, not unpleasant and (often) usable in the field

The remaining figures show examples of the use of simple tests in the field of epidemiological studies on chronic respiratory disease made by my colleagues at the Pneumoconiosis Research Unit (Carpenter *et al*, 1956 Higgins, Oldham Cochrane and Gilson, 1956) Figure 7 demonstrates the importance of the selection of population on the results—a point rarely considered in any detail in most lung function studies We were attempting to find the relation of ventilatory capacity, measured by the indirect maximum voluntary ventilation method to the X ray category of simple pneumoconiosis among coalworkers One group was the working population of a colliery (about 250 men) and the other 300 miners and ex miners applying for compen sation at a pneumoconiosis panel The figure shows the con trary trends in the two groups, a conflict of evidence which showed the need to study a truly representative sample of miners and ex miners

Figure 8 shows the results of an inquiry into the common belief that elderly miners are unduly breathless and subject to bronchitis even though they have not got pneumoconiosis It shows the distribution of maximum voluntary ventilation (in direct) in representative samples of miners ex miners and non miners, aged 55–64 in Leigh in Lancashire The miners have a

attempting to assess a loss of faculty of 10 or even 1 per cent, as now required in this country under the Industrial Injuries Act Figure 6 bears on this problem and shows the relation of a clinical assessment of dyspnoea using a system of four grades

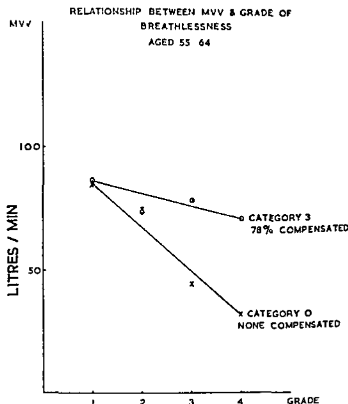


FIG 6 The relation of maximum voluntary ventilation (indirect) 1/min related to a clinical grading of breathlessness in two groups of miners and ex miners aged 55–64 One group (46) with category 3 simple pneumoconiosis the other (47) with no radiological evidence of pneumoconiosis

and the maximum voluntary ventilation (indirect) in a random sample of miners and ex miners, aged 55–64 The results for the miners not receiving compensation for pneumoconiosis make sense, whereas the results for the group who had simple pneumoconiosis, category 3 (nearly 80 per cent of whom were

compensated), make one wonder whether their dyspnoea estimated clinically was not conditioned more by compensation than by lung dysfunction

EPIDEMIOLOGY

Lastly we come to uses for these tests in epidemiology This section should, of course, really come first for it concerns aetiology and prevention, but it happens to be the field in which they have been least used so far

In epidemiology of chest disease we are concerned with the effect of heredity and environment on function We are thus concerned with the results for groups of people and individual variations are less important, but other difficulties arise If the refusal rate in the groups studied is to be kept acceptably low the tests must be short, not unpleasant and (often) usable in the field

The remaining figures show examples of the use of simple tests in the field of epidemiological studies on chronic respiratory disease made by my colleagues at the Pneumoconiosis Research Unit (Carpenter *et al*, 1956 Higgins, Oldham Cochrane and Gilson, 1956) Figure 7 demonstrates the importance of the selection of population on the results—a point rarely considered in any detail in most lung function studies We were attempting to find the relation of ventilatory capacity, measured by the indirect maximum voluntary ventilation method to the X ray category of simple pneumoconiosis among coalworkers One group was the working population of a colliery (about 250 men) and the other 300 miners and ex miners applying for compensation at a pneumoconiosis panel The figure shows the contrary trends in the two groups, a conflict of evidence which showed the need to study a truly representative sample of miners and ex miners

Figure 8 shows the results of an inquiry into the common belief that elderly miners are unduly breathless and subject to bronchitis even though they have not got pneumoconiosis It shows the distribution of maximum voluntary ventilation (indirect) in representative samples of miners, ex miners and non miners, aged 55–64 in Leigh in Lancashire The miners have a

lower average ventilatory capacity and there is a larger propor tion with a maximum voluntary ventilation low enough to account for their breathlessness

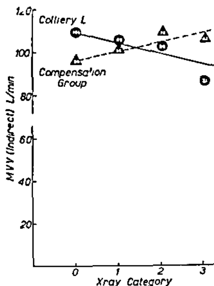


Fig 7 Relation of radiological category of simple pneumoconiosis to age standardized maximum voluntary ventilation in Colliery L and the compen sation groups (M V V s are standardized to age 45)

Table 6 refers to the curious condition encountered especially in the card room workers in the cotton mills—known as byssi nosis—which is at the present being investigated by Dr R S F Schilling Those affected have a tightness of the chest, especially on Mondays or after a period away from work and this is their only symptom in the early stages Later they become progres sively more short of breath and eventually totally incapacitated by respiratory disability An effect on lung function which seems to be specifically related to the mill dust can be detected during the day (McKerrow, McDermott, Gilson and Schilling 1958) The table shows the rise in air flow and tissue resistance measured by an interrupter method during the course of a Monday

This is an unusual condition in which there is both an acute and chronic effect on lung function which can be detected by

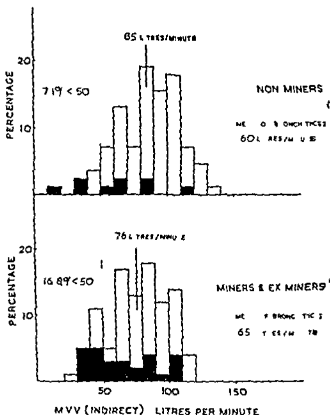


FIG 8 Frequency distributions of M V V Men with chronic bronchitis are indicated by black columns

TABLE 6 Effect of cotton dust during a Monday on airflow and tissue resistance (interrupter method) cm H_2O/l /sec

Age group	No	Resistance 7 30 a m	4 p m	Mean Diff
15–34	7	3 70	6 11	+57 9
35–54	13	3 48	3 87	+17 8
55+	7	4 13	5 56	+32 8
Total	27			+32 1

Resistances > 10 units taken as 10

G

suitable tests We hope to use the acute effect to measure bio logically the improvement produced by dust suppression now being installed in the card rooms

These are only a few examples of the use of these function tests in the epidemiological study of chest diseases Because of the paucity of other objective indices in most cases of chronic respiratory disease, tests of lung function have, I believe, a great future usefulness in this field

SUMMARY

During the last five years tests of lung function have greatly in creased our understanding of pulmonary ventilation and gas exchange But important aspects of lung function in man are still obscure, including the mechanism of the commonest symptom—dyspnoea

Tests have been of more use in assessing the effect of medical and surgical treatment than in diagnosis in pathological terms But a system of functional diagnosis is being developed

In the study of ageing and in the epidemiology of chronic respiratory disease tests have a special usefulness

Lastly, let us take heart that we are some way ahead from the time when Samuel Pepys wrote in his diary in 1665—'But what, among other fine discourse, pleased me most was Sir G Ent about respiration that it is not to this day known, or concluded on among physicians nor to be done either, how the action is managed by nature or for what use it is

REFERENCES

ALBRITTON E C (1953) *Handbook of biological data* 1st fascicule Standard values in blood W D Saunders Philadelphia

ARMITAGE G H and TAYLOR A BRIAN (1956) *Thorax* **11** 281

ASMUSSEN E and NIELSEN M (1950) *J appl Physiol* **3** 95

ATTINGER E O GOLDSTEIN M M and SEGAL M S (1956) *Amer Rev Tuberc* **74** 220

BATES D V BOUCOT N G and DORMER A E (1955) *J Physiol* **129** 237

BATES D V and PEARCE J F (1956) *J Physiol* **132** 232

BEDELL G N MARSHALL R DUBOIS A B and COMROE J H (1956) *J clin Invest* **35** 664
BERNSTEIN L D SILVA J L and MENDEL D (1952) *Thorax* **7** 255
BIERMAN H R KELLY K H CORDES F L PATRAKIS N L KASS H and SHPIL E L (1952) *Blood* **7** 533
CARPENTER R G COCHRANE A L GILSON J C and HIGGINS I T T (1956) *Brit J industr Med* **13** 166
COHN J E CARROLL D G ARMSTRONG B W SHEPARD R H and RILEY R L (1954) *J appl Physiol* **6** 588
COLLIER C R and AFFELDT J E (1954) *J appl Physiol* **6** 531
DAVIES D F and SHOCK N W (1950) *J clin Invest* **29** 496
DAVY H (1839) *Researches Chemical and Philosophical* Smith Elder & Co London
DODDS E C (1953) *Lectures on the Scientific Basis of Medicine* Vol I 1951–52 p 285 The Athlone Press London
DONALD K W (1953) *Brit med J* **i** 1068
DOUGLAS G C (1954) *Handbook of Respiratory Physiology* U S A F School of Aviation Medicine Randolph Air Force Base Texas
DUBOIS A B BOTELHO S Y and COMROE J H (1956) *J clin Invest* **35** 327
DUBOIS A B BOTELHO S Y BEDELL G N MARSHALL R and COMROE J H (1956) *J clin Invest* **35** 322
FILLEY G F MACINTOSH D J and WRIGHT G W A (1954) *J clin Invest* **33** 530
FLETCHER W M and HOPKINS F G (1907) *J Physiol* **35** 247
FRY D L EBERT R V STEAD W W and BROWN C C (1954) *Amer J Med* **16** 80
GAENSLER E A (1955) *New Engl J Med* **252** 177 221, 264
GANDEVIA B and HUGH JONES P (1957) *Thorax* **12** 290
GILSON J C HUGH JONES P OLDHAM P D and MEADE F (1955) Spec Rep Ser med Res Coun (Lond) no 290
HAMILTON M PICKERING G W ROBERTS J A FRASER and SOWRY G S C (1954) *Clin Sci* **13** 11
HIGGINS I T T OLDHAM P D COCHRANE A L and GILSON J C (1956) *Brit med J* **ii** 904
HIRDES J J and BOSCH M W (1955) *J thorac Surg* **30** 719
KENNEDY M C S (1953) *Thorax* **8** 73
KROGH M (1915) *J Physiol* **49** 71
LEUALLEN E C and FOWLER W S (1955) *Amer Rev Tuberc* **72** 783
MCKERROW C B (1955) M D Thesis University of Cambridge
MCKERROW C B MCDERMOTT M GILSON J C and SCHILLING R S F (1958) *Brit J industr Med* **15** 75
MARSHALL R and DUBOIS A B (1956) *Clin Sci* **15** 473
MARSHALL R MCILROY M B and CHRISTIE R V (1954) *Clin Sci* **13** 137

suitable tests We hope to use the acute effect to measure biologically the improvement produced by dust suppression now being installed in the card rooms

These are only a few examples of the use of these function tests in the epidemiological study of chest diseases Because of the paucity of other objective indices in most cases of chronic respiratory disease, tests of lung function have, I believe, a great future usefulness in this field

SUMMARY

During the last five years tests of lung function have greatly increased our understanding of pulmonary ventilation and gas exchange But important aspects of lung function in man are still obscure, including the mechanism of the commonest symptom—dyspnoea

Tests have been of more use in assessing the effect of medical and surgical treatment than in diagnosis in pathological terms But a system of functional diagnosis is being developed

In the study of ageing and in the epidemiology of chronic respiratory disease tests have a special usefulness

Lastly, let us take heart that we are some way ahead from the time when Samuel Pepys wrote in his diary in 1665— But what, among other fine discourse, pleased me most, was Sir G Ent about respiration that it is not to this day known, or concluded on among physicians, nor to be done either, how the action is managed by nature or for what use it is

REFERENCES

ALBRITTON E C (1953) *Handbook of biological data* 1st fasicule Standard values in blood W D Saunders Philadelphia

ARMITAGE G H and TAYLOR A BRIAN (1956) *Thorax* **11** 281

ASMUSSEN E and NIELSEN M (1950) *J appl Physiol* 3 95

ATTINGER E O GOLDSTEIN M M and SEGAL M S (1956) *Amer Rev Tuberc* **74** 220

BATES D V BOUCOT N G and DORMER A E (1955) *J Physiol* **129** 237

BATES D V and PEARCE J F (1956) *J Physiol* **132** 232

VI

The Principles of Heart-Lung Machines

D G MELROSE

A LECTURE on the principles of heart lung machines had it been delivered but a short while ago must have had an interest largely academic Today, however, the very real advances in this subject have translated the main interest to the field of clinical surgery Operations are being performed within the chambers of the heart in several centres throughout the world because machines have been devised which are fully capable of carrying on the functions of the heart and lungs for short periods While by no means yet perfected these machines have established a new approach to the treatment of heart disease and have allowed surgical correction of many cardiac disorders It is particularly difficult at this time to pass judgement on the merits of the several solutions offered for so rapidly is the field developing that anything said today must of necessity be soon supplanted by further advances but certain broad principles can be declared and it is these that I want to discuss

OXYGENATORS

Imitation of the function of the lungs remains the most difficult problem and it is here that opinion is most sharply divided In all four solutions are offered, three of which are wholly artificial, and one involves substituting for the normal lung a lung taken from another animal The homologous lung is a very efficient mechanism for the oxygenation of blood and most cer tainly represents a nearly physiological method for introducing

MEAD J and WHITTENBERGER J L (1952–3) *J appl Physiol* 5 779
NEEDHAM C D ROGAN M C and MCDONALD I (1954) *Thorax* 9 313
PRICE H L CONNER E H and DRIPPS R D (1954) *J appl Physiol* 6 517
RILEY R L and COURNAND A (1949) *J appl Physiol* 1 825
ROHRER F (1915) *Pflug Arch ges Physiol* **162** 225
WEST, J B FOWLER K T, HUGH JONES P and O DONNELL T V (1957) *Clin Sci* **16** 529
WRIGHT B M (1954) *J Physiol* **127** 25P

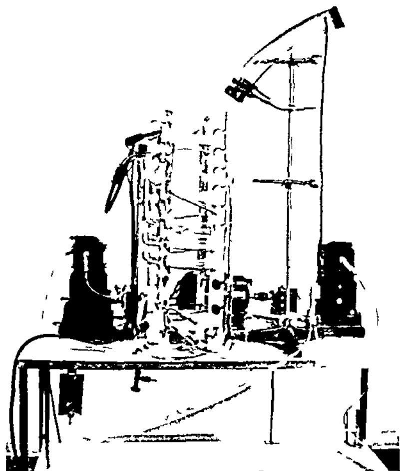


FIG 1 The bubble oxygenator of De Wall with two Sigma motor pumps The debubbling chamber and settling helix will be noted

oxygen into the blood It is, however, not readily prepared and does not lend itself to the maintenance of good sterile technique An alternative to the isolated homologous lung has been suggested whereby a portion of the animal s own lung is used, but cannulation of the pulmonary veins is exceedingly difficult and little success yet attends this attractive possibility

Of the wholly artificial systems, the one which most nearly approaches physiological conditions is that involving the use of semi permeable membranes This technique has evolved from experiences with the artificial kidney However, the membranes available scarcely allow an adequate gas exchange and their use is limited Those most useful at present are made of ethyl cellulose or polyethylene and the most efficient transmission of oxygen through such membranes to date was obtained with an ethyl cellulose film one thousandth of an inch in thickness 14 6 cubic centimetres of oxygen per square metre per minute diffused through such a membrane When using film so thin considerable mechanical problems have to be overcome in the design of an oxygenator based on this principle I have no doubt that the plastics industry could solve all these problems and this technique must deserve considerable further investigation (Clowes *et al*, 1956, Kolff *et al* 1956)

The two methods most commonly chosen are those which involve exposure of blood directly to oxygen In the first of these blood is filmed on bubbles of oxygen In some designs these bubbles are microscopic and a dense foam is created (Clark, Gollan and Gupta 1950) In others bubbles are larger, the foam produced is much less dense, and less difficulty attends reconstruction of blood free from gas bubbles (Clowes, 1954) A large variety of designs to accomplish this form of oxygenation have been described, one of which is used extensively in clinical practice This is the oxygenator ascribed to De Wall and Lillehei (De Wall *et al* 1956), and shown in Plate XII, Figure 1 In it large bubble are formed as oxygen is dispersed in blood and these pass up a vertical tube of polyvinyl chloride At the top of this tube is a debubbling chamber containing a silicone anti foam compound from which the defoamed blood is allowed to descend in a spiral of wide bore tubing This helix acts as a

settling chamber and also as a reservoir from which oxygenated blood can be pumped The tubing is disposable and the unit constructed anew for each perfusion

A more recent modification of this, also disposable, is a single polyvinyl chloride unit in the form of a plastic bag (Hyman 1956) The simplest description of this is that of a pair of plastic trousers, up one leg of which the bubbles climb and after de foaming at the top, cascade down the other leg as blood It is likely that a great variety of fully disposable units will be produced for this is undoubtedly the simplest method of introducing oxygen into blood However, it is as yet not known what physico chemical alterations to the blood such bubbling devices cause, nor whether any such changes are in fact important

In the hands of such master surgeons as Lillehei and Varco at the University of Minnesota in Minneapolis, the bubble oxygenator has allowed at least a hundred intracardiac operations to be performed with a mortality of between twenty and twenty five per cent This, of course, is as yet a frightening percentage However, it has been repeatedly demonstrated in surgery that a new technique even in the hands of the finest exponents carries an alarming mortality, and if this consideration is placed against the inevitable early death of a large proportion of these patients it represents a very significant advance In other hands this device has not allowed all the many problems of cardiac surgery to be solved and if I may venture an opinion it is that this system has certain limitations and is best used in circumstances when perfusion rates are low and operation times are short Improvements are in train and time will show whether these overcome fully the present disadvantages

The bubble oxygenator has in the main been used in association with what is known as the low flow principle This principle was derived from the experiments of Andreasen and Watson (1953) who showed that provided the azygos vein was unobstructed occlusion of both the superior vena cava and inferior vena cava could be maintained for periods of up to thirty minutes without the subsequent death of an animal The blood flow carried by the azygos vein allowed a cardiac output

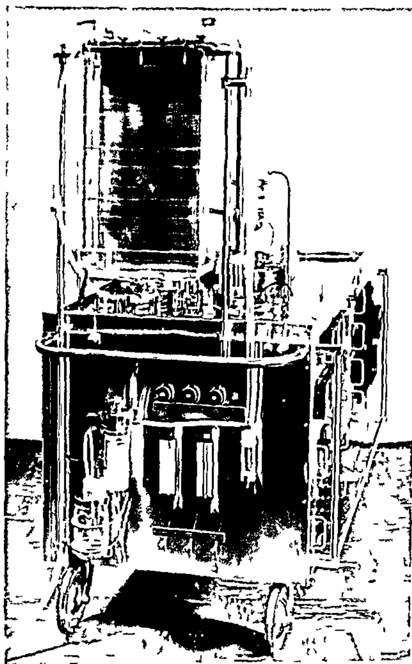


Fig 2 The wire screen oxygenator of the Gibbon apparatus in use at the Mayo Clinic

liquid interface, the surface area exposed being continually renewed as the discs rotate Their device was a great deal more efficient than any previously described and did much to renew interest in this type of oxygenator

J H Gibbon Jr (1937), who published a paper on the artificial maintenance of the circulation during experimental occlusion of the pulmonary artery and who must be regarded as the doyen of workers in this field, substituted for the concentric vertical cylinders which he had previously used a number of wire gauze screens over which blood is streamed These screens are stationary and, as blood descends over them, the turbulence of their passage exposes a great number of cells to oxygen In order to make more efficient such a system which has of itself no moving parts the blood is recirculated within its own pulmonary circuit

It is this method of oxygenation which has been so brilliantly utilized by the workers at the Mayo Clinic Using Gibbon s machine as a model they have constructed an artificial heart and lung capable of circulating and adequately oxygenating over five litres of blood per minute without serious destruction of its elements (Plate XIII, Figure 2) Two disadvantages attend the use of the vertical screen principle One is the difficulty involved in creating a uniform film of blood over the screens There is a tendency for formation of rivulets which of course immediately limits the area of blood exposed The screens themselves cannot be allowed to dry while filming is in progress and hence the film once established cannot be broken without danger of failure to reconstitute it Thus the oxygenator once charged must be kept running throughout what may be a long waiting period The second difficulty is occasioned by the fact that the faster the blood flows over the screens the more blood is in fact held on the screens themselves Hence the blood volume of the artificial lung tends to increase as the flow rate is increased and in order to control this a flow rate in excess of any expected during perfusion must first be established through the pulmonary circuit and then maintained thereafter These disadvantages are only practical ones and should be easily eradicated in the future New materials may themselves bring

of only about ten per cent of the normal, but this low blood flow was in fact sufficient These experiments led Cohen and Lillehei (1954) to use perfusion rates of between twenty and thirty per cent of the expected cardiac output in their human perfusions

This low flow in combination with the bubble oxygenator undoubtedly has been *demonstrated to be effective*, though certain penalties are attached to it A severe metabolic acidosis is created by this inadequate perfusion from which recovery is usual But better recoveries follow a more adequate perfusion and I consider the ideal to be a full replacement and imitation of normal circumstances The primary object of the by pass procedure is to allow surgeons adequate time to repair defects within the heart and if it is a penalty of the low flow principle that this time is limited because of alterations to the acid base balance it must surely be but a stop gap procedure

That the guiding principle should be the full replacement of the circulation with adequate reserves of oxygenation and flow rate is the firm conviction of the group of surgeons, physiologists, anaesthetists and others at the Mayo Clinic (Kirklin *et al*, 1955) This team have demonstrated a remarkable mastery of all the problems involved They have chosen a method of oxygenation which involves the direct exposure of blood to gas but which does not involve the bubbling of gas through blood This principle whereby blood is spread in very thin films and exposed directly is one of the oldest known To enumerate the many methods described to ensure the provision of large surface areas would involve too lengthy an historical review Three examples will be enough to make clear the principles of such a method Frey and Gruber (1885) in their artificial lung allowed blood to spread in a thin film over the inner surface of a cylinder which was filled with oxygen The surface area was about half a square metre Variants of this were tried in succeeding years but proved ineffective Bjork (1948) in Stockholm described a machine using rotating discs to expose films of blood to oxygen In their machine forty or fifty discs dip into a trough of blood When rotated they pick up on their surfaces thin films of blood and in this way create a large gas

liquid interface, the surface area exposed being continually renewed as the discs rotate Their device was a great deal more efficient than any previously described and did much to renew interest in this type of oxygenator

J H Gibbon, Jr (1937) who published a paper on the artificial maintenance of the circulation during experimental occlusion of the pulmonary artery and who must be regarded as the doyen of workers in this field substituted for the concentric vertical cylinders which he had previously used a number of wire gauze screens over which blood is streamed These screens are stationary and, as blood descends over them, the turbulence of their passage exposes a great number of cells to oxygen In order to make more efficient such a system which has of itself no moving parts, the blood is recirculated within its own pulmonary circuit'

It is this method of oxygenation which has been so brilliantly utilized by the workers at the Mayo Clinic Using Gibbon s machine as a model they have constructed an artificial heart and lung capable of circulating and adequately oxygenating over five litres of blood per minute without serious destruction of its elements (Plate XIII, Figure 2) Two disadvantages attend the use of the vertical screen principle One is the difficulty involved in creating a uniform film of blood over the screens There is a tendency for formation of rivulets which of course immediately limits the area of blood exposed The screens themselves cannot be allowed to dry while filming is in progress and hence the film once established cannot be broken without danger of failure to reconstitute it Thus the oxygenator once charged must be kept running throughout what may be a long waiting period The second difficulty is occasioned by the fact that the faster the blood flows over the screens the more blood is in fact held on the screens themselves Hence the blood volume of the artificial lung tends to increase as the flow rate is increased and in order to control this a flow rate in excess of any expected during perfusion must first be established through the pulmonary circuit and then maintained thereafter These disadvantages are only practical ones and should be easily eradicated in the future New materials may themselves bring

relief from these problems and eliminate the present disadvantages

When this problem was taken up at the Postgraduate Medical School of London in 1949 it was decided that the rotating disc oxygenator should be used Experience with the original machine indicated its effectiveness and after comparison with other methods it was decided to develop this principle Simple enlargement was rejected and a novel design adopted In this machine blood passes along a rotating cylinder set at twenty degrees to the horizontal (Plate XIV, Figure 3) In the cylinder thin plastic discs are so arranged that they form crescentic protrusions into the lumen of the cylinder As the cylinder rotates these discs pass under the blood and then up into the gas mixture where the adherent blood films are oxygenated When rotating at a hundred revolutions a minute the available surface area exposed is in the region of a hundred and twenty square metres and oxygen can be introduced into the blood at rates of more than one hundred cc s per litre per minute Destruction of the cellular elements of blood is slight, as is alteration of the blood chemistry

As yet no full comparison of the methods discussed has been possible It is only now that technical progress in the actual application of these machines has reached a sufficient degree of consistency to warrant such a comparison It is likely that in the near future a real reassessment of the position will be undertaken Not only must the actual efficiencies and effectiveness of each of these methods be taken into account but a study of their practicality will often reject an excellent solution on grounds of complexity, expense difficulty in maintenance or cleaning While fully disposable systems have the most appeal they may also be prohibitively expensive It is sometimes better to clean thoroughly a well made item than to risk possible defects in manufacture in a replaceable one

Pumping Circuits

If the number of varieties of oxygenators is large it is greatly exceeded by that of the pumps offered However, these pumps can be divided into two main types On the one hand there are

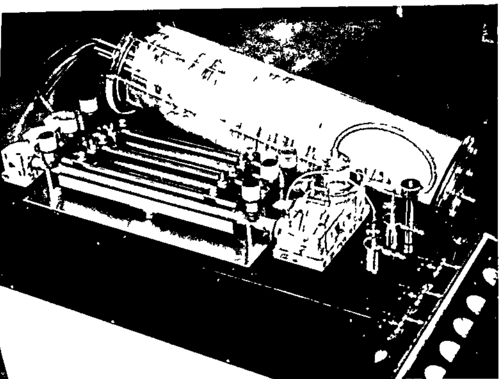


FIG 3 The Melrose rotating disc oxygenator lying on its inclined rollers behind three straight tube pumps

those pumps whose action is to intermittently empty the pump ing chamber the direction of flow being controlled by valves and on the other hand there are those in which an elastic tube through which the blood passes is squeezed within the pump either by the action of a roller or by compressing fingers The essential difference between these two types is that in the first instance control of output of the pump may be both by changing the rate of the pump stroke and also by altering the length of the stroke, while in the second type only a change in rate at which the pump works is possible The pump described by Dale and Schuster (1928) is a good model of the first type while the roller pump of Jouvelet (1934) or that of De Bakey (1934) typifies the second group The Sigma motor pump is essentially a roller pump and should be regarded as such

In use, of course the difference between these two types be comes obvious The intermittent pumps employing valves to control the direction of flow can be made to imitate closely the normal shape of the pulse, while the roller type of pump tends to have a continuous output and a minimum pulse com ponent

The most widely used pump today is that manufactured as the Sigma motor pump Just as the bubble oxygenator is particularly suitable for the low flow type of perfusion so is this form of pump most effective at low flows It is convenient inex pensive but limited in output Attempts to obtain high flows result in considerable destruction of cells

The Gibbon apparatus at the Mayo Clinic uses the roller pump of De Bakey and no difficulty is found in pumping at flow rates in excess of five litres per minute Both these pumps produce an almost continuous flow with minimum pulse, and while no definite evidence exists as to whether a pulsatile flow is necessary it may be important and a pump has been designed which seems closer to the ideal (Melrose 1955)

In this a simple straight rubber or plastic tube is squeezed by compression plates one plate acting as a back stop against which the other plate presses the tube The pumping plate is driven by a cam shaft which imparts an undulating movement to it The back stop can be moved towards or away from this moving

plate and hence the amount of compression can be varied and with it the stroke volume In practice a variation of output between zero and five litres per minute can be obtained in this manner Such a pump of course requires valves to direct the flow and in order to maintain the advantages of the roller pump in this respect these valves do not lie within the blood stream They are essentially miniature versions of the pumping plates and completely occlude the tubing at appropriate points in the pumping cycle which provides a close imitation of the normal ventricular cycle In this design the tubing is only com pressed completely at the valve sites and hence there is little danger of blood destruction by the apposition of the walls of the tube in the pump chamber It is this apposition of the tube walls throughout its whole length in the roller type of pump which endangers the blood cells

It will be seen that no hard and fast rules yet exist to govern the design of pumps Again, the considerations of practical utility have to be balanced against the claims of physiological imitation One thing above all is clear, however, and that is that the pump must be capable of maintaining a blood flow of at least four litres a minute against normal vascular resistance, it is also essential that as smooth a flow as is possible through the pump should be established, that the material with which blood comes in contact should be inert and non toxic that the pump chamber should be capable of sterilization that com plete mechanical reliability can be expected Simplicity, ease of maintenance and accuracy of control of the output from the pump are also valuable attributes The choice is a wide one and there is little save experience to take as a guide

The somewhat bewildering variety of oxygenators and pumps at present being successfully used must indicate how fluid this subject is at present The restless search for the ideal apparatus indicates too that none of the solutions so far offered is wholly satisfactory, and it is true to say that this subject has entered a new phase Now designers can turn their attention to the per fection of these instruments certain that the technique of by passing both heart and lungs is no pipe dream Where before this endeavour was directed to the proving that such a

substitution could be undertaken, now it must be directed to the search for those essential requirements which will enable blood to be circulated extra corporeally with complete safety

The By pass Procedure

It has been the experience of all those working in this field, whatever the solution chosen, that suddenly their work is crowned with success Each group has found as technique builds up and their experience widens that perfusions become possible and problems previously insurmountable melt away The mechanical aspects of the heart lung machine are only, of course but a small part of the technique of the by pass procedure which must become the responsibility of a team experienced not only in the surgery of the heart but also skilled in the measurements of vascular pressures and of biochemical and haematological changes

Nothing categorical can, of course, be laid down in answer to all these ancillary problems However, certain broad principles are agreed at this time to govern the technique of connecting heart lung machines to the patient or animal, to prevent blood coagulation and its attendant problem of re establishing a normal clotting time after perfusion, and to avoid severe metabolic alterations

Cannulation

It is generally agreed that venous blood should be extracted from the great veins entering the heart by wide bore tubes and in practice it is usual that these tubes be passed into the superior and inferior vena cavae through the right atrial appendage This is particularly necessary when blood is allowed to flow freely from these veins without the assistance of a suction pump As yet there is no complete agreement on the vexed problem of how to ensure the maximum flow from the veins, some rely on gravity others on direct suction, but all are agreed that as large a calibre tube as possible should be inserted close to the entrance of the great veins into the heart

For the return of oxygenated blood into the arterial side the

most usual site is the left or right subclavian artery The subclavian artery i of large bore, is close to the vital centres and can be sacrificed if need be without danger Into this vessel can be inserted a cannula capable of carrying the normal cardiac out put for that particular patient without the imposition of a high resistance to the inflow of blood It is essential that a variety of cannulae are at hand to ensure that the largest that will fit into any given subclavian artery can be inserted, for the interference to blood flow created by a very narrow cannula causes con siderable destruction of red cells and seriously hinders per fusion

PRESSURE MEASUREMENT

The measurement of the arterial pressure and also that within either the superior or inferior vena cava by direct cannulation should be regarded as obligatory That obtaining in the great veins is in my opinion the mo t important single factor in establishing a successful perfusion The rate at which the perfusion can be performed must of course depend wholly on the rate at which blood can be obtained from the venous system Additional suction from the veins can have no effect other than to collapse them and to bring the perfusion to a halt It is thus essential to maintain that perfusion rate which maintains the central venous pressure as normally as possible and it is most useful to have a continuous measurement of this value both prior, during and after perfusion I have mentioned the warning given by a reduction of the venous pressure to low levels, herald ing as it does collapse of the veins Equally useful is a rise in venous pressure for its indication that the perfusion rate is not being maintained at the highest possible level Here, of course, the subject gets more controversial for in the low flow system a high venous pressure is expected because the perfusion rate is arbitrarily about a third of the normal cardiac output

A record of the arterial pressure when taken in conjunction with that of the central venous pressure gives a reliable index of the state of the heart and the onset of heart failure can be readily appreciated The arterial pressure is used also to detect changes in the size of the vascular bed and reflect the vaso motor tone

Electrocardiogram

Two other quantities which should if possible be measured continuously are the electrocardiogram and the electroence phalogram Though little change is normally found in the electrocardiogram alteration of the T wave is a sensitive index to chemical change in the blood and of course the onset of arrhythmias gives warning of the possibility of ventricular fibrillation The recording of the electrocardiogram may well be continued for many hours post operatively for the extensive operative repair of some cardiac defects gives rise to consider able anxiety during the period of operative recovery

Detailed interpretation of the electroencephalogram is not a necessity but alterations of the normal alpha and theta rhythms, particularly towards the slow delta rhythm signify a failure to sufficiently perfuse and oxygenate the brain As patients are normally only lightly anaesthetized during cardiac surgery the effects of anaesthetic agents are minimal and alterations to the electroencephalographic pattern are of serious import It has been the experience of many that even an apparently satis factory by pass procedure may come to nought due to an un detected cortical damage if this measurement is neglected

Blood Volume

Very considerable care must be exercised in regard to the blood volume Accurate estimations of blood loss must be made throughout the procedure and full replacements immediately made Care must be exercised to avoid the consequences of the injection of large quantities of fluid, which though isotonic may in fact not be iso osmotic A higher proportion of failure parti- cularly in those cases already incapacitated by pulmonary damage is likely if severe haemodilution occurs So seriously do the workers at the Mayo Clinic consider this problem that they take special care by the addition of human albumen to preserve the iso osmotic value of any fluid used, either to main tain an intravenous route or to flush through cannulae or catheters used for measurement The smaller the patient the more important this aspect is and the more essential it is to

preserve a normal blood volume Weighing before and after perfusion should be carried out and readjustments made to ensure no overloading of the vascular system

Anticoagulants

All machines require a certain amount of blood to charge them This blood must be fresh blood taken through non wettable tubing into non wettable containers The anticoagulant used should in general be heparin and a generally agreed value for this is fifteen milligrammes of heparin to each five hundred millilitres of blood stored Coagulation is prevented in the patient or animal by the intravenous injection of heparin Alternatives exist, of course, but heparin is by far the most common substance used Though in some respects inadequate, particularly in regard to the change in the actual physical properties of blood which heparin creates, heparin has the merit of being a naturally occurring substance whose action can be reversed This reversal can be brought about by the injection of protamine sulphate This compound forms a loose bond with free heparin in the blood and neutralizes its anticoagulant effect

Protamine is not without dangers and of itself creates clotting problems if given in great excess It has an additional disadvantage of effecting the blood pressure if given rapidly or in too high a concentration A workable rule is to obtain a heparin level in excess of two milligrammes per kilogram of body weight in the blood and to reverse it at the end of perfusion with the same quantity of protamine after which a clotting time examination is made If this value is less than twice normal no further protamine need be given If it is greater than this then protamine should be given until twice the estimated dose has been administered

Failure to re establish a normal coagulation time following a by pass procedure may be due to many complicated effects There can be no doubt that platelets are usually deposited on the artificial material of which the apparatus is constructed, and it is known that a form of defibrination of the blood frequently takes place This is not obvious and does not often lead to frank clotting However it does lead to the consumption of

many clotting factors and the only way in which this can be prevented is to ensure that the heparin level is not allowed to fall unduly during perfusion The treatment of the bleeding tendency sometimes found is at present confined to replacement of blood, and in some instances to the addition of fibrinogen Blood must be quite fresh and must be drawn with the same precautions as that used to charge the machine, namely through non wetting surfaces into non wetting containers A direct transfusion can be of enormous value in a severe bleeding tendency A point to note in this context is that if the blood is allowed to cool during passage through the machine the tendency to bleed will of course be exaggerated This problem merits considerable further study for it is by no means solved and continues to jeopardize the by pass technique

Materials

The effect of passing blood over foreign surfaces is a variable one and many of the difficulties associated with coagulation deformities following perfusion are in fact due to the use of inappropriate materials The advent of many new plastic materials whose surface can be rendered quite non wettable and whose composition is non toxic has greatly simplified the choice Platelets, leucocytes and inevitably fibrin, will adhere to rough surfaces to those contaminated by debris and particularly to those exerting strong chemical attraction In this respect untreated glassware is suspect as is rubber in which there is an excess of chemical fillers Pure latex, and glass whose surface has been rendered inactive by a coating of silicone is less guilty in this respect Probably the finest material from which to make apparatus at present is polytetrafluorethylene This has as the trade names fluon or teflon Unfortunately it is as yet, very expensive and most workers are content with either perspex or polyvinyl chloride These two though almost ideal, cannot be safely autoclaved and while chemical sterilization is undoubtedly effective there are many who would prefer to add to such a method the more traditional one of autoclaving

Whatever the material the surface to which blood is to be

exposed must be as smooth as possible It has been found that provided the finest polish is given to it stainless steel is an excel lent material, and is particularly suitable where rigid parts are required The choice of materials grows ever larger and it is quite safe to say that the precise requirements will in future be met

The use of silicone compounds has been particularly reward ing Not only will certain of these compounds make inert other wise active surfaces, but others of the same group have the remarkable ability to de foam blood without in any way injuring the blood itself It would be true to say that the resurgence of interest in the bubble type of oxygenator is wholly dependent upon the discovery of 'Silicone antifoam A

Filtration

The necessity to filter all the blood coming from the machine has not been established However, normally the lungs shield the cerebral and coronary vessels and probably filter out aggre gations of protein and other cellular debris This can be imi tated by screens of fine mesh, but if to be fully effective against such particles these screens constitute a gross impediment to flow and may of themselves contribute largely to the destruc tion along the tubes Most workers are content to screen only those particles which will not pass through relatively large meshes and it is usual to employ a mesh of between a hundred and two hundred and fifty microns In order that hold up by such a filter should not be excessive large areas of mesh must be used Frequently the filter can be combined with an air trap in the circuit

Conclusion

These then are some of the associated problems with their pre sent solutions There are many more particularly those which I will call the physiological response to the by pass procedure Little as yet is known of the physiological mechanism brought into play when the heart and lungs are by passed The be haviour of individual organs is as yet, unstudied, and even the effect on the cellular elements of the blood itself awaits more refined techniques of determining cell survival and function

But a point in development has undoubtedly been reached from which one can look forward to an increasing concentration on such problems rather than those of mere survival

I, myself, am wholly optimistic My wish would be for a fully artificial solution employing a disposable membrane type of oxygenator and pump capable of accurately reproducing the normal pulse This device would require but a small quantity of blood to fill it, be autoclavable and have an oxygenating and pumping capacity closely approaching that of a normal man I think I can speak for most of those who have worked longest in this field when I say that we have just begun to solve the problem posed by the heart lung machine

REFERENCES

ANDREASEN A. T and WATSON F (1953) *Brit J Surg* **41** 195

BJORK V O (1948) *Acta chir scand Supp* **137**

CLARK L E GOLLAN F and GUPTA V B (1950) *Science* **111** 85

CLOWES G H A. Jr (1954) *Surgery* **36** 557

CLOWES G H A. Jr HOPKINS A L and NEVILLE W E (1956) *J thor Surg* **32** 630

COHEN M and LILLEHEI C W (1954) *Surg Gynec Obst* **98** 225

DALE H H and SCHUSTER E H (1928) *J Physiol* **64** 356

DE BAKEY M E (1934) *New Orleans Med Surg* **87** 366

DE WALL R A. WARDEN H E READ R C GOTT V L ZIEGLER N R. VARCO R L and LILLEHEI C W (1956) *Surg Clinics North America* **36** 1025

FREY and GRUBER (1885) *Arch f Anat Physiol* **9** 519

GIBBON J H Jr (1937) *Arch Surg* **34** 1105

HYMAN E S (1956) *Trans Am Soc Art Int Organs* **ii** 1

JOUVELET — (1934) *Bull Soc Méd Hop Paris* **50** 537

KIRKLIN J W DUSHANE J W PATRICK R T DONALD D E HETZEL P S HARSHBARGER H G and WOOD E H (1955) *Proc Mayo Clinic* **30** 201

KOLFF W J EFFLER D B GROVES L J PEEREBOOM G and MORACA P P (1956) *Cleveland Clin Quart* **23** 69

MELROSE D G (1955) *J Physiol* **127** 51P

VII

Cajal and Sherrington

E G T LIDDELL

Histology 1800–65

WHETHER or not the task of the medical biographer is a weighty one depends in part upon what is generally known about the subject. It is general knowledge, for instance, that before Lister's discovery of antisepsis, the mortality of surgical cases was terribly high. The history of wars and of hospitals makes that clear. So when that admirable biography of Lord Lister was composed by Sir Rickman Godlee, there existed already a generally appreciated background of information against which the reader could gauge Lister's achievement. But when it comes to describing the secluded life of laboratory workers the task is not quite so simple, because what work is attempted and what is done in laboratories is not so generally known or hardly known at all. The background of information is scanty, gauging the achievement difficult. Both Cajal and Sherrington were laboratory workers. Cajal was born in 1852 and this year of 1957 is Sherrington's centenary year. Although they never worked together their ideas were linked together in the last decade of the century to support and succour the so called Neurone Theory—the idea that nerve cells in the nervous system were separate entities—each unit separated anatomically but not functionally from every other unit. What came before this Neurone Theory? What had people thought? What had they tried to do? Why was this theory so different? To get the fullest measure of the background, I will go back to the beginning of the nineteenth century.

At that time, here in London, we should all have been very

much aware of a colourful, not to say flamboyant character, Sir Everard Home He wrote much on microscopical observations and was well familar with early microscopes, because he had borrowed (but alas! never returned) the valuable collection of Leeuwenhoek's microscopes which belonged to the Royal Society For the present purpose I just wish to record his warning which is always true, that we should never trust too much the picture made by the microscope This was especially true in those early days because the microscopes had every possible defect, beginning with chromatic aberration It was not until 1826 that the father to be of Lord Lister, Joseph Jackson Lister a London business man caused the first achromatic microscope to be made In 1829 the Royal Society published his paper upon it The two dates mark the beginning of the end of the stone age of microscopy Before that time nerve fibres floated in water as was the practice, developed regular varicosities, which, when viewed with defective microscopes, gave the appearance of a string of beads—globular bodies, they were called This created confusion when a little later, globular bodies which were, in fact, nerve cells, began to be described

Perhaps the last worker in that early time who merits a word is the botanist Dutrochet the author of an attractive little monograph (1824) on the appearance and excitability of vegetable and animal tissues Among his illustrations one finds the picture of a peripheral nerve fibre floating in water with little bodies perched here and there upon its surface like buns on a plate which he said were nerve cells And he described similar bodies closely packed together in the brain of the frog which he said were identical in appearance and were also nerve cells No doubt he was looking at myelin globules and not at nerve cells *at all* certainly not on the surface of a peripheral nerve fibre But he did say that the brain is an organ eminently destined for the production of nerve force since the preponderant tissue is nerve corpuscles He showed a drawing of nerve fibres from some of which a granular substance is emerging since pressure has been applied to the coverslip This showed at any rate, that nerve fibres were not hollow tubes, as had been thought some years earlier Therefore, if they were not hollow

tubes, it would be difficult for 'animal spirits', 'nerve force', 'vital force', to travel through them So the idea of animal spirits, inherited from Galen by Willis, Descartes and other writers, was henceforward to be doubted Mention of pressure on the coverslip brings me to the technical methods of those times There had been no fixatives, though alcohol was beginning to come into use Hitherto, in order to obtain a translucent section of the central nervous system, the microscopist had to cut, with curved scissors or a 'very sharp knife, as best he could at the surface of a tissue which resembled a jellified soup in which globular objects might be vaguely seen He would then press on the coverslip so as to make the section more translucent Peripheral nerve fibres were much tougher to handle but difficult to see into It seems strange that alcohol was not much used before 1811, remembering that Boyle had described its preservative powers, that Settala's anatomical museum existed in Milan at the end of the seventeenth century with its bottled Siamese twins and was one of the sights of the city, and that the British Navy was reputed to have knowledge of the preservation of the bodies of deceased admirals by immersion in a cask of rum There were no good fixatives, no microtomes, no serial sections no stains But until achromatic microscopes were available, the need for these improvements was not apparent As Dutrochet said in his time, and as Everard Home before him, the illusions of the microscope make it difficult to distinguish the truth

The first credit for describing nerve cells is usually given to Ehrenberg (1833) who saw and described in a very brief account, irregular globules in the brain, and also 'ganglionic globules or corpuscles in the spinal ganglia of frogs His microscope is not mentioned as being achromatic, but worked best at a three hundred fold magnification In the following year Rudolph Wagner mentioned that the nerve fibres of sight, smell and hearing organs suddenly swell into bladder shaped bodies *full of nerve* marrow, after which they narrow down again into fine nerve tubes which are crystal clear Nerve tubes elsewhere, however, were filled with nerve marrow

The first continental author to sing the praises of the achro

matic microscope was Purkinje, a Czech born professor at Breslau, in Germany He had acquired an instrument from Vienna in the early thirties, and using it 'with real wolf s hunger described various ganglionic corpuscles in parts of the brain and spinal cord (Ehrenberg s name of ganglionic cor puscles' was to continue to be used for a long time) Purkinje also used a microtome but has left no description of it It is probable that he also, as early as 1830, had een the nucleus in the yolk of the hen's egg, a year before Brown saw the nucleus as an opaque spot' in the vegetable cell Purkinje s great day came on 23 September 1837 when he described to a meeting of scientists and doctors at Prague those cells in the cerebellum which are still called after him 'Each of these small round bodies has a blunt rounded end on the inner side towards the yellow substance and shows in its retort shaped form a central nucleus The other tail like end is directed outwards and is lost mostly with two processes going into the outer grey substance where the vascular layer is Here is an advance to be noted Nerve cells are described and clearly illustrated They have nuclei and they have processes It is a proper tribute that these cerebellar cells have always been known as Purkinje cells

Remak s monograph which is so often quoted did not appear until the following year, 1838 and with illustrations which are less good, of nerve cells in sympathetic and dorsal root ganglia, with nuclei inside the cell and processes protruding from the surface of the cell In 1841 Remak went on further to describe the pale protoplasmic processes which afterwards came to be called dendrites Before long thanks to the labours of one of Purkinje s pupils Valentin the central nervous cor puscle was being likened to an ovum with the nucleus like the germ the contents of the outer enveloping cell like the yolk, and the wall of the cell to the vitelline membrane Thus the earliest days were passing There was no longer any reason for observers to peer with simple magnifying glasses or no glasses at all at imperfectly prepared material and pronounce that the central grey matter was cineritious that is grey, amorphous ash like But still in the decade before Cajal and Sherrington were born regular methods were hardly established and results were

tubes, it would be difficult for 'animal spirits', 'nerve force', 'vital force', to travel through them So the idea of animal spirits, inherited from Galen by Willis, Descartes and other writers, was henceforward to be doubted Mention of pressure on the coverslip brings me to the technical methods of those times There had been no fixatives, though alcohol was beginning to come into use Hitherto, in order to obtain a translucent section of the central nervous system, the microscopist had to cut, with curved scissors or a 'very sharp knife', as best he could, at the surface of a tissue which resembled a jellified soup in which globular objects might be vaguely seen He would then press on the coverslip so as to make the section more translucent Peripheral nerve fibres were much tougher to handle but difficult to see into It seems strange that alcohol was not much used before 1811, remembering that Boyle had described its preservative powers, that Settala's anatomical museum existed in Milan at the end of the seventeenth century with its bottled Siamese twins and was one of the sights of the city, and that the British Navy was reputed to have knowledge of the preservation of the bodies of deceased admirals by immersion in a cask of rum There were no good fixatives, no microtomes, no serial sections, no stains But until achromatic microscopes were available, the need for these improvements was not apparent As Dutrochet said in his time, and as Everard Home before him, 'the illusions of the microscope make it difficult to distinguish the truth'

The first credit for describing nerve cells is usually given to Ehrenberg (1833) who saw and described in a very brief account, irregular globules in the brain, and also 'ganglionic globules' or 'corpuscles in the spinal ganglia of frogs His microscope is not mentioned as being achromatic, but worked best at a three hundred fold magnification In the following year Rudolph Wagner mentioned that the nerve fibres of sight, smell and hearing organs suddenly swell into bladder shaped bodies full of nerve marrow, after which they narrow down again into fine nerve tubes which are crystal clear Nerve tubes elsewhere, however, were filled with nerve marrow

The first continental author to sing the praises of the achro

When I saw the magnificent radiating bundles and the cortical tracts of nerves I realized that I had discovered the key which would open the chambers of the wonderful structure of the spinal cord Not more gladly did Archimedes call out his *ευρεκα* than did I shout out on beholding that sight' Even if 'that sight wasn't as wonderful as sights seen in later years Stilling had at least introduced the notion of freezing before sectioning, and reaffirmed the notion of serial section

An early form of microtome consisted of a tube with a flanged mouth across which the very sharp knife' was drawn by hand to cut each section while a thumbscrew in the tube pushed the material towards the mouth Tissues were not embedded as yet, but were shaped into a cylinder by a sort of cork borer and pushed into the microtome tube from the flanged mouth If that procedure was not possible they were encased with pith or turnip so as to make a tight fit inside the tube When freezing microtomes did come the evaporation of ether was used to produce the freeze Much later on in 1869, Edwin Klebs made a very casual reference to paraffin wax for embedding and in 1871 Rutherford of Edinburgh dropped paraffin wax into the microtome tube in order to consolidate the position of the tissue in its orifice The idea of using paraffin seems to have grown up quickly and silently from no particular origin, and at the end of the century Gustav Mann in his learned book on histological technique confessed that he did not know who introduced the method of embedding in paraffin wax

At mid century the outstanding book on histology was by Kolliker, translated from German into English in 1853 by George Busk curator of the musuem at the Royal College of Surgeons and Thomas Huxley Kolliker ever cautious and in his long life usually right (he lived from 1817 to 1905) collated the existing evidence regarding ganglionic globules or nerve cells with processes which seemed to pass over into dark bordered tubes That passing over was now a crucial question in the complex higher animals Did all the long processes of nerve cells really always pass into medullated nerves? Technique was not yet good enough to allow the process from the nerve cell to be followed as far as the point where it acquired a

often very much a matter of chance At least, however, about this time, the general improvement in living conditions which followed the construction of railways had its good effect also upon science, for communication between places and the transport of books and papers were much facilitated Scientists could exchange information more easily

In 1842 three events took place which were important for miscroscopical science First, Helmholtz, who, of course, began his career as an army surgeon and was then twenty one, wrote his M D thesis for the University of Berlin Working, it is true, on invertebrates, with the microscope uncorrected for chromatic abberation, he described clearly how he saw the processes from nerve cells running into the primitive tubes, as medullated nerve fibres were then called Nerve cells and nerve fibres at last had been discovered to be associated with one another 'Et caudas eorum in ipsos nervos transeuntes The long caudal process of nerve cells ran into a nerve fibre But other nerve fibres merely came from central parts' Second, there came a thesis from a Dane named Hannover who, like Helmholtz, had worked in the Berlin laboratory of Johannes Muller, and made the important introduction of potassium bichromate as a fixative Besides other advantages its use got rid, once and for all, of varicosities in nerve fibres—or rather prevented their appearance Hannover s monograph, published in Copenhagen, is remarkable for some of its illustrations, probably the best that had hitherto appeared Moreover, some of his material was from human subjects His nerve cells with their processes are easily recognizable for what they are Before then, the reader would need to look at the text, before the figure to find out what he was supposed to see The length of the processes from nerve cells were better shown for a distance equal to four or five diameters but after that they were lost Stains were yet to come, serial sections were unreliable, but came soon as the third important event of 1842 On 25 January 1842 Stilling by chance left a piece of spinal cord on the window sill at a temperature of -13 R The next morning, finding it frozen hard he cut sections and examined them under his magnifying glass a poor affair with a magnification of only fifteen diameters or less

definite and emphatic statements about the separateness of nerve cells Their processes do not join with or anastomose with the processes of any other nerve cells They terminate freely or approach the neighbourhood of other nerve cells The cells are entirely independent units Anyone who thinks that the processes of nerve cells join or anastomose believes in a delusion Deiters in fact had discovered the Neurone Theory but his early death prevented recognition in a time that was not ripe The part of his work that was recognized was the sure establishment of the belief that every nerve fibre orginated in a nerve cell, and was never a separate kind of tissue This came to be known as Deiters Law Twenty years before some authors had confessed that nerve fibres were not to be distinguished from fibrous tissue That time had now passed Deiters Law was accepted, but not seemingly with outspoken enthusiasm in its own country of origin Later on, in this country E A Schafer supported it strongly The slow growth of realization of scientific truth in those times could have been noted in the previous decade on a different problem Augustus Waller, after describing in 1850 his famous observations on the alterations produced in the structure of the fibres of the glosspharyngeal and hypoglossal nerves of the frog, following section went on two years later to experiment with the nerves on either side of the dorsal root ganglion and concluded that the nutritive centre for the fibres lay in the ganglion—just that, and no more He made no suggestions about the nature of the nutritive centre Nevertheless, his degenerate nerves attracted much attention, as they deserved to do Soon everyone was talking about Wallerian degeneration without knowing the real reason for degeneration

Golgi and Cajal the Whole Structure of the Nerve Cell With Deiters nearly forgotten after never being fully recognized the year 1873 was memorable for the introduction by Camillo Golgi of Pavia of his method of staining nerve cells Silver as a reagent was twenty years old but Golgi s method of applying it, essentially a simple method gave startling results The nerve cell and all its processes stood out as one whole black silhouette with a clarity that had never been seen before and which is still

sheath So in 1858 Claude Bernard, a bold experimentalist eager to advance and not an over cautious histologist inhibited by lack of evidence, pointed out the obvious by saying that nerve tubes must have both a beginning and an end Besides nerve tubes, he said, there was another kind of element in the nervous system, namely nerve cells which were found in the grey matter of the cord and in ganglia Nerve cells he continued, were sometimes the origin of nerve fibres Indeed in the anterior column of grey matter in the spinal cord there is a condition of large ganglionic corpuscles which give rise to the motor nerves Stilling had established so much in 1842, and had seen nerve fibres coming out of nerve cells in the oculo-motor and hypoglossal nuclei Some histologists about this time thought that the connection between nerve cells and nerve fibres might be only temporary, and that nerve fibres, after being secreted by nerve cells were separated off from them

Deiters a Forgotten Histologist

An answer to these doubts came in 1865 but outrageous fortune decreed that the answer should not be widely heard The answer was contained in a monograph published two years after the author's death from typhus at the age of twenty nine The author was Deiters and the editor of the posthumous book was his teacher Max Schultze of Bonn Deiters had used potassium bichromate as a fixative he cut his sections in series and he used a stain, ammoniacal carmine That was better, he found than the ordinary carmine which was the red ink of the period and still exists in our kitchens today as cochineal Besides introducing these three strong points of technique, Deiters also made some excellent drawings There can be no doubt that he saw clearly and distinctly as he said clearly and emphatically, that every nerve cell gives rise to a long process which becomes a nerve fibre That long process looks different and is different He even noted the existence of local narrowing of the process soon after it leaves the body of its parent nerve cell The other processes from the nerve cell, the clear protoplasmic processes, 'dendrites as they were called later, were altogether different from that long process Most importantly Deiters made most

their heretical opinions angered the Spaniard to his dying day One unvanquished Reticularist was Dogiel of St Petersburg who, *using Ehrlich s new methylene blue stain claimed to see*, at least in the sympathetic nervous system that there were proto plasmic bridges across points of anastomosis of nerve cell processes By 1891, however, the modern view was firmly set on the map by Waldeyer of Berlin who suggested that this new theory of separate nerve cell units should be called the Neurone Theory Waldeyer's only contribution had been to sit and think of a word! Van Gehuchten of Louvain had been a strong supporter of Cajal His objection was to the use of the word Theory', because facts to him were facts , not theories A fresh develop ment of the idea was started by van Gehuchten who in a vague tenuous way made suggestions which led Cajal on to describe his notion of the 'dynamic polarization of nerve cells Cajal s argument was that the cells gave origin to the special structure of the cylindraxis with its characteristic property of conducting nerve messages away from the cell body to a special destination The dendrites, the clear protoplasmic prolongations of the cell were concered with centripetal conduction towards the body of the cell Even two years earlier Cajal had formed the mental picture of dendritic prolongations being collectors or receptors of electrical currents Cajal did not use the word polarization in an electrical but in a geographical sense To the nerve cell there was a way in as it were and there was a way out The beginning pole had its own function and the outgoing pole had its own function Perhaps , Cajal wrote in later years a new technique in the future may reveal new and more intimate connections between nerve cells At the present, it has to be admitted that the nerve currents are transmitted from one element to another by a sort of induction or influence at a distance Cajal s mention of induction or influence at a dis *tance between nerve cells was of course* an old idea which had been put forward by E A Schafer as early as 1878 Cajal does not seem to have known that Schafer had seen the separated nerve cells in the umbrella of *Aurelia* intermingling intertwin ing though without anastomosis or joining Schafer had said then that the conveyance of nerve message between nerve cells must

very familiar to us all Golgi himself lost no time in publishing many impressive illustrations but he adhered to the current belief, undisturbed as it had been by Deiters, that a nerve cell by its long processes anastomosed end to end with other nerve cells by their long processes and so were protoplasmically continuous with one another in a diffuse network The finer processes (dendrites) were supposed to be for nourishment, like the roots of a tree Many histologists began to use the Golgi method, among them August Forel of Zurich As Forel says in his autobiography, he was using Golgi's method in the autumn of 1886 and began then to have doubts of Golgi s network The nerve cell processes seemed to him to end blindly, without anastomosis He came to be quite sure that the idea of Golgi's diffuse network of anastomosing processes resulted from erroneous observation Thinking over his own results and their meaning, *he remembered Ranvier s work which showed how processes* grow out of nerve cells again after injury Gudden's atrophy method, too, showed that only certain groups of cells showed change when their axones had been cut The changes were not general Then, only in October of that year, His had shown how protoplasmic processes grow out of embryonic nerve cells All this was evidence of the separateness of nerve cells from one another, and in January 1887 Forel published his doubts But to his later regret, he did not think of a name for the new theory for which his results gave support By 1889 Ramon y Cajal in Barcelona, who had been using the Golgi method for a year on embryonic tissues, was also quite sure that 'Reticularists who believed in the Golgi network, were wrong He attended a meeting in Berlin in the first fortnight of October of that year and showed his preparations Important authorities were there, Kolliker, the doyen of histology in Germany and the world, Retzius from Stockholm, His from Leipzig Waldeyer from Berlin These world celebrities began their examination with more scepticism than curiosity Undoubtedly they expected a fiasco, wrote the sensitive Cajal, but he goes on Finally, the prejudice against the humble Spanish anatomist vanished and warm and sincere congratulations burst forth There and then Reticularists were almost entirely routed But

SHERRINGTON THE FUNCTION OF THE NEURONE

Sherrington was five and a half years younger than Cajal having been born in 1857 As an assistant to J N Langley he began in 1884 to publish papers with him on the spinal and other degenerations produced by Goltz's excisions from the brains of dogs—all this let it be remembered, before the days of Forel, Cajal and Marchi, when without great risk of con tradiction it was still possible to believe—as many did—in the Golgi diffuse reticulum The best stain was still carmine, but there was no sure knowledge of the relation whether direct or indirect, between the cortex cerebri and the pyramidal tract Then in a few years time when Sherrington was in his early thirties came the revelations of Forel Cajal and van Gehuchten In the experimental fields in which Sherrington became so lucid an exponent, it had long been clear enough that the Bell Majendie law of spinal conduction was unshakeably true in the main The law of forward direction (William James, 1890) when extrapolated to the single cell became the law of dynamic polarization No matter how much the motor spinal root was stimulated (except in such a way as to give gross escape of electrical current) messages did not travel from its origin back into the central nervous system so far as could be seen at that time The circuits in the nervous system appeared to Sherring ton to be valved against regurgitation The nerve centre exhibits a valve like function, allowing conduction to occur through it in one direction only That the direction of nerve impulses is not reversible along the neural chains may be the function of the synapse If at the nexus between neurone and neurone there does not exist any actual confluence there must be a surface of separation Such a surface might restrain diffu sion bank up osmotic pressure, restrict the movement of ions, accumulate electric changes support a double electric layer, alter in shape and surface tension with changes in difference of potential alter in difference of potential with changes in sur face tension or in shape or intervene as a membrane between dilute solutions of electrolytes of different concentration or colloidal suspensions with different sign of charge The latest

be by 'inductive action whether electrical or of some other kind Or perhaps exuberant speculation was being restricted at that time by Kolliker s warning that what was seen to exist in one kind of animal must not be assumed to exist in all kinds of animals

The final touch to the Neurone Theory' came in 1897 when Held described 'end feet' upon the bodies of nerve cells They were the blunt ends of axons from other nerve cells Held did not use a silver stain, and it remained for Cajal to produce an improved silver stain which showed clearly the bulbous processes ending blindly on the surface of nerve cells and their dendrites Contiguity, not continuity', was the watchword of the neuronists These end feet', 'end knobs', boutons terminaux, are very familiar to modern scientists and provide many mysterious problems for them at the present time Meanwhile, in the seventh edition of Foster s *Physiology* in which the named author was assisted by Sherrington, there appears the word synapse And in Schafer s textbook of 1900 where Sherrington is again an author, the word appears once more It is convenient to have a short term by which to speak of the union of one nerve cell with another by this close contiguity of processes with the cell body or of processes with processes presumably without actual continuity of substance Sherrington describes the nerve cells as holding hands with one another That belief in contiguity and not anastomosis was to be the working hypothesis of himself and most other experimentalists for years to come, but doubters continued to exist sporadically Golgi s last public confession of faith as a Reticularist was made at Stockholm in 1906 when he and Cajal shared the Nobel Prize The ceremony was a dramatic occasion for the older man Himself the inventor of the silver staining method, he declared himself unable to agree with the findings of Cajal who owed discoveries mainly to that method Yet both were Laureates together Golgi said that he maintained his belief in the diffuse nerve network between cells, but he did concede that the neurone was an independent cellular unit

professor The old idea remained in books until 1891 when it appeared in Stirling s textbook, but with a correcting caption to say that owing to the discoveries of Cayal (*sic*) the diagram must now be considered out of date In its most modern and perfected form however it reappeared in 1925 to illustrate the phenomenon of recruitment of motoneurones in reflex excitation (Sherrington, 1925 *Proc Roy Soc* B 97, 519) In the nineties the idea ripened in Sherrington s mind that Foster s lines of resistance in the nerve net were the 'intercellular barriers, deli cate transverse membranes existing at the synapses where nerve cells 'held hands They were also the valves against re gurgitation So at the turn of the century, when he had settled the wearisome geographical problem of limb plexuses, Sherrington felt ready to find out in detail what characters dis tinguished reflex arc conduction from nerve trunk conduction One experimental feature of his work was different from others He made graphic records of muscular action, records that could be preserved and analysed Many of his predecessors had used myography for studying the activities of peripheral nerve-muscle preparations but not often of reflex actions François Franck had been early in the field to record myograms when the cerebral cortex was stimulated Sherrington bought his copy of Franck s book in 1887 the year of its publication, and had those beautiful illustrations at his elbow—beautiful be cause their author was a worthy pupil of Marey the early master of graphic recording who eventually gave up physio logy to become a pioneer of research in motion pictures Among physiologists of an earlier generation who had not made graphic records was Ludwig (1853) but he had made the preg nant observation that a stimulus thrown into the spinal nervous system might suffer temporal and qualitative alterations and might elicit a repetitive response Some years later, Gergens in Goltz s laboratory had noted and written on the scratch reflex of spinal mammals which was a variation of the acidu lated paper experiment on the spinal frog used when Pfluger was arguing in favour of the *Ruckenmarkseele* (spinal cord soul) So Sherrington by picking up these old suggestions and making careful documentation of muscular acts in the scratch reflex

researches with the electron microscope show clearly the surface of separation between a bouton and the dendrite of the next cell in the reflex chain, so that the wealth of possibilities suggested by Sherrington sixty years ago would not be seriously alien from modern thought

Some specificity of canalization through the central nervous system had long been a notion just out of range Experimentalists of the nineteenth century had been full of enthusiasm but had garnered many tares with their grain Results were plentiful but so was muddlement At mid century in 1853 Carl Ludwig had suggested and shown in his textbook a diagram of central nerve paths in which nerve fibres formed a network with nerve cells at the nodes Ingoing nerve messages entered along one strand, and, after passing a node, arrived at a muscle fibre If the nerve messages were stronger, they went past a second nerve cell in its node, along another outgoing strand to another muscle fibre Still stronger messages went still further through the network, past further nodes, and reached still further muscle fibres Ganglionic bodies in the cord determined the direction of excitation of pathways This simple reticular scheme of excitatory propagation had been envisaged by Rudolph Wagner as early as the 1830s It continued to be illustrated or referred to in textbooks Michael Foster in the 1879 edition (third) of his book wrote We may infer that the protoplasmic network spoken of above is, so to speak, mapped out into nervous mechanisms by the establishment of lines of greater or less resistance, so that the disturbances in it generated by certain afferent impulses are directed into certain efferent channels But the arrangement of these mechanisms is not a fixed and rigid one It is possible to suppose that the lines of resistance in the spinal protoplasm are so arranged as to admit of an alternative action were there any evidence of a variable automatism like that of conscious volition being manifested by the spinal cord of the frog we should be justified in supposing that the choice was determined by an intelligence ' It is interesting to note that the third edition of Michael Foster's book was current when Sherrington went as a freshman to Cambridge, and would convey to him the opinion of the period and of his

REFERENCES

To seek out early original references may be tedious and frustrating The following books however give accounts of neurophysiology in its early days Some of them are well furnished with references to original papers

Books

CAJAL S RAMON Y *Recollections of My Life* Trans E Horne Craigie Philadelphia 1937

DEITERS O F K *Untersuchungen uber Gehirn und Ruckenmark des Menschen und der Saugetiere* Braunschweig 1865

FOREL A *Ruckblick auf mein Leben* Zurich 1935

VAN GEHUCHTEN A *Anatomie du systeme nerveux de l homme* Louvain 1806

HANNOVER A H *Mikroskopisker Undersogelser af Nervesystemet* Copenhagen 1843 French trans Paris and Leipzig 1844

KOLLIKER A *Manual of Human Histology* vol I Trans by G Busk and T Huxley London 1853

LONGET F A *Anatomie et physiologie du système nerveux de l homme et des animaux* 2 vols Paris 1842

LONGET F A *Traite de physiologie* 2nd edn (2 vols) Paris 1860–1 3rd edn (3 vols) Paris 1868–9

MANN G *Physiological Histology* Oxford 1902

SCHAFER E A *Textbook of Physiology* vol II Edinburgh and London 1900

Original Paper

SCHAFER E A (1878) *Philos Trans* 169 563

as well as by the selection of fruitful questions, established an impregnable array of evidence for the differences between re actions in the spinal nervous system and reaction in peripheral nerve The latency of response, the after discharge long after the electric flea had stopped biting, the facilitation of one weak stimulus by another weak stimulus travelling from a different segment of the spinal cord, the regularity of rhythm whatever the rate of the 'bite of the electric flea, the long refractory period in each beat when many bites fell in an unresponsive phase, the fatiguability of this grooming scratch reflex as contrasted with the relative non fatiguability of a prepotent danger avoiding reflex like the flexor reflex, the inhibition of the scratch reflex by the flexor reflex, its rebound with wide swinging strokes after inhibition, its irreversibility, were all special features of the scratch reflex and resulted from the passage of the nerve impulse through spinal reflex centres The scratch reflex was, and per haps still is, the clearest all round illustration of the character istics of reflex action Nerve axones when studied peripherally behaved quite differently The difference of behaviour be tween axones and nerve centres must surely lie in the nerve centres themselves in those synapses where nerve cells touched, but did not join with other nerve cells Sherrington by this analysis of reflex action established his reputation as a master But he could not have achieved his results without careful graphic analysis of reflex action To observe the action only without making permanent measurable records would have missed most of the most critical points Knowledge would have lagged and not have stepped forward as it did when Sherrington made known his Observations on the scratch reflex in the spinal dog' Although usually thought of as an experimenter, Sher rington had also high standards in microscopy His early days had been spent in bacteriology and he had been a pupil of J N Langley Structure for him came first, then function First, Cajal, the anatomist, and then Sherrington, the physio logist, used in their separate studies methods which were good And the questions asked of Nature were good, too, and had great answers

REFERENCES

To seek out early original references may be tedious and frustrating The following books however give accounts of neurophysiology in its early days Some of them are well furnished with references to original papers

Books

CAJAL S RAMON Y *Recollections of My Life* Trans E Horne Craigie Philadelphia 1937

DEITERS O F K *Untersuchungen uber Gehirn und Ruckenmark des Menschen und der Saugetiere* Braunschweig 1865

FOREL A. *Ruckblick auf mein Leben* Zurich 1935

VAN GEHUCHTEN A *Anatomie du système nerveux de l homme* Louvain 1806

HANNOVER A H *Mikroskopisker Undersogelser af Nervesystemet* Copenhagen 1843 French trans Paris and Leipzig 1844

KOLLIKER A *Manual of Human Histology* vol I Trans by G Busk and T Huxley London 1853

LONGET F A. *Anatomie et physiologie du système nerveux de l homme et des animaux* 2 vols Paris 1842

LONGET F A. *Traite de physiologie* 2nd edn (2 vols) Paris 1860–1 3rd edn (3 vols) Paris 1868–9

MANN G *Physiological Histology* Oxford 1902

SCHAFER E A. *Textbook of Physiology* vol II Edinburgh and London 1900

Original Paper

SCHAFER E A. (1878) *Philos Trans* 169 563

VIII

Social Psychiatry

AUBREY LEWIS

PSYCHIATRY is mostly thought of as being concerned with the individual Research has been concentrated on how individual patients behave, the intricate psychological mechanisms in each individual, and the metabolic anomalies or tissue changes that accompany and underlie his disturbed con duct In this, psychiatry has of course been following the main stream of medicine Pathology, whether it be psychopathology or cellular and chemical pathology has been looked for inside the patient s body and mind

Alongside this traditional and necessary approach to the study of mental disorder however there has been a tentative less explicit awareness of extraneous forces In primitive societies, and indeed in Western Europe until a couple of centuries ago this was expressed chiefly in crude animistic superstitions about witchcraft and possession by evil spirits As we became more sophisticated and critical such beliefs were of course given up, but the impact of outside forces upon the mind of each indivi dual was still plain enough The causes of drug addiction pro vide a glaring example

The converse was also plain Every insane person, every un stable neurotic, had an effect upon others which showed his inability to conform, in one way or another, to the customary requirements of ordered society his failure in this regard was one of the main reasons for considering him mentally ill

The social causes of mental disturbance and its social effects, were therefore perceptible to everybody who reflected on the

matter At certain periods of history the evidence was so dramatic that it forced itself even on the unreflecting There were the epidemics of flagellation and of crazy dancing which broke out in many German cities in the twelfth, thirteenth and fourteenth centuries, disrupting the regular life of the community and bespeaking grossly disorganized social conditions A later example is the affair of the Salem witches at the end of the seventeenth century A number of adolescent girls, evidently hysterical, declared that they were being harried by the devil at the instance of various people whom they named In the course of one year (1691–2) two hundred and fifty members of the small New England community where this occurred were arrested and tried twenty of these innocent people were judicially murdered or tortured to death and the colony was distracted by the fear and dissension engendered (Starkey, 1952)

There was another powerful reminder of how social influences worked upon psychiatry A revolution was effected in the care of the insane at the end of the eighteenth century instead of shackling them in cells beating and whipping them into submission, Pinel in France and Tuke in England introduced the humane and compassionate regime which has ever since been the practice in civilized countries The reform was of course a manifest product of the immense social forces then at work which released humanitarian impulses of many kinds

It is hardly necessary to recall the ways in which the study of mental illness must be incomplete if social factors are not taken fully into account (Simmons and Wolff 1944) The truth is, however, that until lately social factors were not studied or even properly allowed for in the clinical assessment of prognosis and treatment of mental illness The main reasons for this comparative neglect were the bias in medical education and the lack of a well developed science of sociology with appropriate techniques concepts and theoretical foundations The claims of psychoanalysis to explain all human behaviour likewise diverted attention from the social causes and effects of mental abnormality All this has changed in the last decade or two and I shall attempt in this lecture to outline the methods of social investigation that have been used to illuminate psychiatry in the last

twenty years I shall begin with the Hutterite study, carried out in 1950 and 1951 (Eaton and Weil, 1955)

The Hutterites

The Hutterites are a Protestant sect that originated in Switzerland early in the sixteenth century From the beginning they were persecuted, and had to flee from one country to another, at first from Moravia, then from Hungary, later from Carinthia and Turkey In 1774 they emigrated to the Ukraine and then, just a century later, moved in a body to the United States A large section of the Hutterites continued to live in colonies which have been virtually closed communities for the last seventy years Only about 2 5 per cent of the colony population have left for the outside world in that period Some of the colonies are in Canada, in Alberta and Manitoba others in Montana and Dakota, in the United States

The religion of the Hutterites forbids marriage outside the sect and they believe in the common ownership of all property The community carries the responsibility for its members, it provides them with their clothing, pocket money and food and looks after them when they are ill They have no wages, no birth control, practically no celibacy (except for some of the mentally disturbed and defective) and no divorce The average size of family is eleven or twelve children The reproduction and survival rates are fairly uniform throughout the population, showing none of the variations with class or educational level which are familiar in other modern populations It is a stable, but far from stationary, society, with a fairly constant distribution of persons by age and sex It tends to double its size every sixteen years

These are almost unique features suggestive of some ideal pastoral community of earlier times, but there are also, in dramatic contrast, many of the familiar features of present-day Western society The Hutterites use tractors and refrigerators and up to date machinery, they avail themselves of the resources of modern medicine, and schooling is on the usual American or Canadian pattern

In this seeming Utopia crime and insanity were said by

credible witnesses to be almost unknown Doctors and sociologists agreed that this was so, and the Hutterites' immunity from mental ills was credited to the freedom from tension and conflict inherent in their way of life—a way of life remote from the financial stresses, family dissension and competitive struggles for prestige and power and enjoyment which are evident in most other societies

Now it was of considerable importance to preventive medicine to determine whether, in fact the mental health of the Hutterite community is so notably good as all these observers reported It would illuminate, it might even settle a long standing controversy about the effect of competitive Western civilization upon the incidence of mental disorder This controversy is typified on the one hand by gloomy prophets who declare that mental disorder is steadily on the increase in our troubled times and on the other hand by psychoanalysts who regard the experiences of early life within the family as far more decisive for mental health than the trials of daily life and the vicissitudes of the social order I shall again touch on this question briefly later

A field survey was therefore made of the Hutterite colonies It revealed that of the entire community alive at the time of the survey (1951) amounting to 8 542 persons fifty three had at some time had a psychotic breakdown sixty nine had had a neurotic illness and fifty one were mentally defective, there were also twenty epileptics and six people with psychopathic personality making 199 in all who had had some psychiatric disability that is about 2 3 per cent of the whole community More than half of these were exhibiting the mental disability at the time of the inquiry and another quarter had recovered by this time

This disposed of the hitherto well attested reputation of the Hutterites for almost complete freedom from mental troubles It was in keeping with a similar finding on the Navaho Indians who had been repeatedly described as stolid and contented but were found on systematic examination to be anxious and emotionally unstable Its moral was that a consensus of informed opinion is not to be trusted in a matter of this kind only

competent surveys will do But the study raised a number of questions which have to be answered in nearly all social and epidemiological investigations, especially psychiatric ones

Was the method of ascertainment trustworthy and fully effective? Is it permissible to lump together diverse conditions—in this case psychoses and neuroses and mental defect? If not, how reliable are the diagnoses which serve to designate the separate classes of illness investigated? And can one compare the attack rates and prevalence rates in one community with those in another whose social characteristics are also known, and so draw conclusions about the relative effect of social variables?

It was only after these questions had been answered that it was found that although the Hutterites were not collectively blessed with the perfect mental health attributed to them, they were better off in this respect than most others of whom we have detailed knowledge, and that sociological factors are indeed significant in promoting mental well being

But this broad conclusion has to be qualified for example it does not apply equally well to psychoses and to neurotic disorders as conventionally distinguished It is necessary to look at the question more closely

First, the method of ascertainment It is not uncommon to take admission to a mental hospital as the criterion of mental disorder, or more strictly, of psychotic disorder But the number of people who go into a mental hospital depends, in appreciable measure, on the number of beds available, the administrative arrangements and the attitude of doctors and the public towards these hospitals—which may be feared or trusted The ascertainment of neuroses is a still more difficult undertaking since there is room for much difference of opinion about marginal cases—which are numerous—and the fullness of the information available may greatly affect the count Thus in the Hutterite study the colonies were dealt with in three ways—nineteen of them were investigated intensively sixty five were investigated less closely with brief interviews and the remaining nine colonies were assessed on the basis of information provided by doctors and by members of the other colonies, including preachers who travelled about to the colonies in question

In the intensively studied population (1,671 persons), 13 8 per 1 000 were judged neurotic, in the less intensively studied (6,123 persons), 7 2 per 1 000 were so diagnosed and in the population studied through information supplied by others, 2 7 per 1 000 It is most unlikely that these striking differences reflect any real difference in incidence of neuroses

This problem of ascertainment is the most serious that waylays the investigator of comparative incidence of mental disorder It is obviously crucial Compared with the epidemiological study of physical diseases like pulmonary tuberculosis, psychiatric work in this field is beset with diagnostic uncertainties To overcome them, there has to be much preliminary work This is aimed at gaining the confidence of the people studied, so that they will give full information, and at developing uniform, operational criteria of a case which will be applied consistently by all those engaged in the work of ascertainment Though the difficulties inherent in this do not preclude one from comparing the findings in one study say in Scandinavia with those in another say in America or England they do enforce caution in doing so In the Hutterite study, for instance the amount of psychotic illness found would give an expectancy of nearly three times as much psychosis among them as among the rural inhabitants of Thuringia and Bavaria who had been studied earlier by a German psychiatrist (Brugger, 1931, 1938) But scrutiny of the German investigations indicates that patients who had recovered or were for other reasons outside a mental hospital at the time of the inquiry were likely to have been missed and the diagnostic criteria used were evidently different In regard to neuroses firm operational definitions have to be formulated and adhered to steadily This is even more necessary with psychopathic personality—a notoriously unsatisfactory diagnostic category (Lewis 1953) Anti social conduct however (which of course may be the outcome and evidence of psychopathic personality) is a less ambiguous term in a given community, and its amount can be measured Among the 8 542 *Hutterites* it was practically unknown There were none who had committed major crimes such as murder or violent assault or had been involved in a sexual offence only

twelve had been found guilty of theft, usually on only one occasion, and there was no juvenile delinquency (apart from one youth who had been trapping animals for their fur without a licence). Alchoholism and other drug addiction was almost non existent.

The Hutterites, then—that idyllic community—have slightly less psychotic illness than other rural groups in Europe and North America (though depressive disorder is in higher proportion than usual, as against schizophrenia). They have probably the same amount, or perhaps less, neurotic and related psychogenic illness than others do, though this is hard to estimate; they do not engage at all in antisocial activities or violence, and their children do not become disturbed enough to require special action on the part of parents or doctors.

In the light of these findings, should one conclude that the Hutterites are a particularly healthy society? The question is a deceptive one. An individual's mental health can be determined but not that of a society, unless the term 'mental health' is to be given a different meaning. It can, however, be concluded that the Hutterite community is more peaceable and law abiding than most, and that this is achieved at the cost of an austere conformity which restricts inventiveness and enterprise.

From the study of this interesting community, so strikingly different from its neighbours, much knowledge was expected, but in the event it cast less light on the effect of cultural and social influences upon mental health than had been hoped for. Eaton and Weil, the very able investigators who carried it out and reported it, sum it up: 'Our findings do not confirm the hypothesis that a simple and relatively uncomplicated way of life provides virtual immunity from mental disorders. ... Psychoses and other forms of mental disorder were found to occur with regularity in the Hutterite population. Their existence in so secure and stable a social order suggests that there are genetic, organic and constitutional elements which predispose a few individuals to mental breakdown in any social system, no matter how protective and well integrated it may be. ... A mental health Utopia is probably impossible.'

Class Incidence

If the close scrutiny of the Hutterites yielded so little evidence which would reflect social and cultural influences upon the amount of mental illness or health, it is perhaps hardly to be supposed that differences in incidence would be found between people more uniformly exposed to one culture But it has frequently been noted that there is a disparity between the urban and rural rates of mental illness and between the rates for different occupations in the same county Thus in 1953 in Canada urban residents entered mental hospitals as first admissions at approximately a four to three rate over rural dwellers in several provinces the urban rate was twice the rural rate (Marshall, 1953) Much of the disparity here depends it is true on the differing accessibility of services for ascertaining and treating mental disorder But such an explanation could hardly account for the much more dramatic discrepancy in admission rates between people of different occupation and social class Thus in England and Wales in the year 1951 admission rates for males in social class 5 were more than twice those for any other class (Registrar General 1955) In Canada the first admission rate for labourers was three times that for professional workers When the individual diseases are looked at the figures are even more striking Schizophrenia shows an immense class difference Whereas 164 men per million of social classes 1 and 2 over 20 years of age were admitted to mental hospitals in England and Wales in 1951 with a diagnosis of schizophrenia for class 5 the corresponding number was 803—about five times as high a rate

This remarkable class difference has been noted elsewhere A classical study bearing on the matter was carried out in 1937 (Faris and Dunham, 1939) upon patients admitted for the first time to psychiatric hospitals, they found the schizophrenics came to a large extent from homes in densely populated areas of low economic status and social disorganization A finding of this kind can be interpreted in two main ways that the conflicts and privations in crowded urban slums favour the development of schizophrenia, or that schizophrenics as they deteriorate drift into these areas

If the latter explanation were correct, the schizophrenics of social class 5 should include many who had been descending in the social scale or who had been born elsewhere This was looked into in another ecological study (Hollingshead and Redlich, 1954) The proportion of schizophrenics who had been born elsewhere and 'drifted' into the area where the diagnosis was made was estimated, it did not confirm the 'drift' hypo thesis Neither did a survey of the social class of the patients' families support the 'deterioration' hypothesis Ninety one per cent of the patients were still in the same social class as their families, and as to those who had moved socially, more had moved into a higher than into a lower social class On these New Haven figures (which are, however, at variance with another American study (Gerard and Huston, 1953)) it could be asserted that drift and social mobility do not account for the predominance of social class 5 among schizophrenics

But before accepting the alternative explanation, viz that *the stresses of poverty stricken urban life promote schizophrenia*, a further social possibility has to be considered It may be that the schizophrenic from social class 5 is more likely to be referred to a mental hospital though his illness is not severe, or it may even be that people with a mental disorder coming from such a background are more likely to be diagnosed as having *schizophrenia* It would lead us rather far to go into these possibilities No studies that have so far been made have settled the issue and the class disparities are so great on the available statistics, that several factors may be at work The subject is now being studied by at least two groups of investigators in this country, as well as by people in the United States, where most of the work has hitherto been done

So far we have considered two kinds of inquiry one comparing the amount of mental disorder in the simple religious people of a rural enclave with that in the competitive communities surrounding them, and the other comparing the amount of *schizophrenic* mental disorder in the lowest social class with that in social classes 1 and 2 Neither of these types of inquiry has given a clear answer to the etiological questions which obviously arise, but both stimulate our curiosity further about the possible social

causes of mental disorder, and both indicate that the problems are complex, that summary conclusions and unconfirmed opinions are misleading and that systematic methods are essential

All the studies of social causation in psychiatry depend on *two requirements adequate detection and measurement of* mental disorder in a specified population and adequate measurement or description of social characteristics in that population Unless these requirements are met systematic study is impossible If they are met, we can make correlation studies within the population, or we can compare populations to see how far and in what way social differences are associated with psychiatric differences It is of course impossible to measure all the psychiatric phenomena that might be relevant or all the social phenomena, investigations therefore deal with selected variables chosen to test a hypothesis The hypotheses have come to a considerable extent from psychiatry and psychoanalysis but some of the most brilliant ideas have come from sociologists like Durkheim, and most stimulating investigations have been made by social ecologists like Faris and Dunham

MARITAL STATUS EXILE EMIGRATION

None of these studies of course, can do much more than point to an association between psychiatric and social variables How the association comes about and in what sequence has to be studied through rather different inquiries, of a more individual and clinical character Examples of this may be seen in studies of marital status and of exile or emigration

Everyone has observed that at every age and in every diagnostic group the hospital admission rate for single persons is appreciably higher than for the married This is especially the case with schizophrenia and manic depressive psychosis and it is much more striking in schizophrenia than in manic depressive psychosis In other words a strong association can be demonstrated between the incidence of major psychoses and the social fact of marriage This can be extended on the social side to determine where the widowed stand in this regard how sex and economic and class status affect the relation how the proportions differ between countries in which early marriage is the

rule or divorce is common, whether the rate among people who are necessarily celibate, like Roman Catholic priests and nuns (Moore, 1936), varies from that of other single people, and so forth But the three main explanations put forward (Odegaard, 1946, Norris, 1956) are (1) that a single person is more readily admitted to a mental hospital than a married person, or (2) that those who develop mental disease present pre morbid personality traits which militate against marriage, so that the married are a positively selected group, or (3) that in married life there are factors which protect against mental diseases Now in order to look for support or rebuttal of these possible explanations, an intensive study of individuals became necessary Ødegaard found that in the single patients before their breakdown, unfavourable features of personality (especially schizoid traits) were twice as common as among the married, and that the sexual outlook of the single patients had been passive and cold, especially in the patients with schizoid personality The second explanation was therefore supported There are other findings of this kind reported in the study, which was conducted in two stages—the first, an epidemiological analysis of first admissions between 1926 and 1939, altogether 14,231 patients, and the second an intensive clinical study of 707 patients personally examined by the investigator

Sainsbury as I shall mention in a moment followed a somewhat similar procedure in studying suicide There is no opposition between the study of individuals (by clinical or psychological methods) and the study of populations—on the contrary, they complement and reinforce one another

The same combination of extensive and intensive methods has been employed in comparing the mental health of immigrants with that of the population they spring from Thus Norwegians who migrated to Minnesota had a higher admission rate to mental hospitals than Norwegians in Norway, this was especially so in respect of schizophrenia To discover whether this should be attributed to the stresses of adapting to a new country or, rather, to the migrants containing a higher proportion of persons predisposed to mental illness a careful study (Odegaard, 1932) was made of the personal history and clinical

state of each psychotic migrant in the sample and then when it appeared that schizoid features, or insidiously developing schizophrenia, had been already evident in the migrants at the time of their leaving Norway a further scrutiny was made of the course and pattern of illness in the schizophrenic immigrants, to see whether the content of their delusions and similar clinical features testified to the stresses of adapting to their new environment This was not the case their clinical picture was in almost all respects identical with that seen in Norwegian mental hospitals The outcome of these detailed extensions of the study was to support the view that factors in the personality which predispose to schizophrenic breakdown may also conduce to making a man decide to emigrate

Broadly we may say that the methods used in socio psychiatric studies thus derive from epidemiology and ecology, from anthropology, sociology and social psychology and from clinical psychiatry

The epidemiological method we have already considered As ascertainment must be effective it cannot depend wholly on *hospital admissions nor on incomplete volunteer populations*, nor on detection of cases by inexpert persons or by doctors with diverse criteria of illness But because findings open to these objections are obviously subject to error, they are not without value as sources of stimulating hypotheses and as pointers towards further more detailed and dependable studies

Incomplete Data

An example of the usefulness of incomplete data is to be seen in an inquiry by Goldhamer and Marshall (1953) It has long been customary to assert that mental disorder is on the increase because of the more and more complex strains which society imposes on its members A hundred and thirty years ago the psychiatrist Esquirol read a paper in Paris on the question *Are there more madmen today than there were forty years ago?* There was then widespread alarm at the frightful increase of insanity which people said menaced France with calamity In his admirably critical review Esquirol showed that the alleged increase was spurious more people had come into

hospitals, it was true, but there had been tremendous advances in the facilities for treatment of mental illness, and the reputation of the asylums had changed from that of cruel prisons to humanely conducted hospitals Many quiet patients who would not formerly have been sent to these institutions were now admitted, and came willingly Hence the apparent in crease Esquirol, in his well reasoned article, disposed of the lament about the increase in insanity—but only for a short time the same cry has often been raised since

Goldhamer and Marshall decided to check the truth of this assertion They surveyed the first admissions to Massachusetts mental hospitals for the hundred years beginning 1840 They found, at first blush, a great increase The crude admission rate per 100,000 at risk in 1840–5 was less than half what it had become in 1940 But when age specific figures were calculated, the increase was seen to be wholly due to the large number of admissions nowadays of persons over the age of fifty For the other age groups there had been no increase, although the hospital facilities were of course more plentiful, the public attitude was more favourable, and the hospitals were now admitting neurotic disorders also The changes that had occurred *in the structure of New England society* and the large Irish immigration had not made any appreciable difference in the first admission rate for people under fifty (though there had been fluctuations in the proportions in which the various dis eases were represented) The increase in admissions over the age of fifty was evidently, as in Esquirol s day, an indication not of increased incidence but of changing social habits, in the last century families were large and accustomed to look after their elderly members even when demented unless their con duct was intolerable We cannot, therefore, attribute an in crease in mental disorder to the conditions of modern society, with all its struggles and insecurity so long as we have no con vincing evidence that there has in fact been any increase But the main point here is that Goldhamer and *Marshall s study* showed how useful conclusions can be drawn from necessarily incomplete data, provided that the limitations of the data are fully recognized

Suicide

Another instance of how incomplete data can be pressed into service is to be seen in Sainsbury s study of suicide (Sainsbury, 1955) We have had, for two thousand years or so assertions about the morality of suicide, the motives that may prompt or justify it, and the ways to stamp it out Then people became more interested in its causes and in the nineteenth century it was well recognized that besides the special circumstances that affected each individual, factors of more general and lasting force were at work As Buckle put it, in 1857 All the evidence we possess respecting it can leave no doubt in our minds that suicide is merely the product of the general condition of society Although there are still many who conceive of every suicide as a response to individual misery and who are dumb founded when no evident cause can be discerned it is now plain to anyone who looks into the statistics that social forces operate to determine the gross amount of suicide in a commu nity What the social forces are and how the victims are selected has however been obscure

To inquire into these problems by intensive clinical study of individuals is made almost impossible by the mournful fact that the man who commits suicide has put himself beyond our reach Those who attempt suicide and survive won t do be cause they differ in some essential respects from those who die (Stengel 1957) Even when we have had ample data about individuals we have lacked the clue that would guide us safely to the social influences, for about the strength and nature of these influences different psychiatrists expressed divergent opinions based on their clinical experience

Sainsbury decided to use the ecological method He assumed that in a city like London it is possible to differentiate neigh bourhoods with particular social characteristics related to the degree of social mobility and social isolation that prevail within them and he proposed to test the hypothesis that where social mobility and social isolation are pronounced community life will be unstable without order or purpose and that this will be reflected to a greater or less degree in the suicide rates because

men and women are more prone to suicide when they live in, but apart from, a social group which neither achnowledges nor provides means for satisfying their needs He therefore correlated suicide rates in the twenty eight metropolitan boroughs and the City of London with selected social indices, and as a more intensive and narrow check, he analysed the social and medical information available about 409 persons whose suicide had been reported to the North London coroner during a three year period

Sainsbury found that the London boroughs differed significantly from one another in their suicide rates, and showed the same rank order consistently over three decades, although there had in that time been considerable changes in the composition of the resident population The suicide rates correlated significantly with the indices of social isolation (living alone or in a boarding house), social mobility (daily turnover of population, and number of immigrants), and divorce and illegitimacy rates The suicide rate tended to increase in the middle class and decrease with poverty and it had no relation to the amount of overcrowding and unemployment in the various boroughs When, however, the individual records were examined, the frequency of unemployment among the suicides was higher than in the general population—no doubt partly because a man s illness prevented him from working and partly because dismissal from work had helped to provoke gloom or despair

For convincing reasons, Sainsbury concluded that a lonely way of life and some social disorganization favour the occurrence of suicide Taken by itself, it is not a startling conclusion, but it must be seen alongside the negative findings—for example that there was no intrinsic connection between overcrowding and suicide though it was previously thought that there is, and that drift of predisposed persons into the areas that had high suicide rates did not account for these rates

An ecological study has serious limitations It presupposes *social homogeneity in the ecological area (Clausen and Kohn,* 1954) It cannot penetrate into the details of the nexus between social variables and mental illness rate It can only in exceptionally favourable circumstances enable us to determine which

social variables in an area are responsible for that area's rate of mental disorder It can, however indicate where a close search into a presumptive causal relationship is well warranted and where it has little prospect of reward It does this best when it deals with a clinical anomaly that can be fairly well defined (like suicide), when it has areas whose relevant social characteristics are distinctive and readily ascertainable, and most important of all, where there is a well founded hypothesis to be tested

In Sainsbury s study the main hypothesis and the method, derived from Emile Durkheim (Durkheim, 1951) Durkheim had divided the social causes of suicide into three classes those which promote excessive individualism and isolation secondly, those which exercise compelling authority over the individual, driving him into conforming or altruistic acts such as suttee or giving one s life for another and thirdly, those which bespeak social disorganization—anomic forces as Durkheim called them, which leave the individual unsupported by the social controls and assurances that should keep him steady and secure This threefold classification by Durkheim has commended itself to many subsequent workers though they have mostly paid more attention to the psychological aspects of the matter than he did

Sainsbury found social isolation to be the predominant social factor in the populations he studied A similar conclusion could be drawn from a more exotic study made in Singapore (Murphy 1954) The population of Singapore falls into five main ethnic groups—Chinese Indian European Eurasian and Malay each with a different suicide rate The Chinese, who make up three quarters of the total population have a high suicide rate (21 per 100 000) and so have the Indians whereas the Malays (like Muslims everywhere) have a low rate—only 1 7 suicides per 100,000 of population Within the Chinese community there are wide variations in suicide rate (from 12 to 24 in men and from 6 to 16 per 100 000 in women) and these variations are associated with tribal factors the Chinese are immigrants from five different regions of China and their place of origin can be recognized from the dialect they speak The smaller and more economically insecure the dialectal sub

men and women are more prone to suicide when they live in, but apart from, a social group which neither acknowledges nor provides means for satisfying their needs' He therefore correlated suicide rates in the twenty eight metropolitan boroughs and the City of London with selected social indices, and, as a more intensive and narrow check, he analysed the social and medical information available about 409 persons whose suicide had been reported to the North London coroner during a three year period

Sainsbury found that the London boroughs differed significantly from one another in their suicide rates, and showed the same rank order consistently over three decades, although there had in that time been considerable changes in the composition of the resident population The suicide rates correlated significantly with the indices of social isolation (living alone or in a boarding house), social mobility (daily turnover of population, and number of immigrants), and divorce and illegitimacy rates The suicide rate tended to increase in the middle class and decrease with poverty, and it had no relation to the amount of overcrowding and unemployment in the various boroughs When, however, the individual records were examined, the frequency of unemployment among the suicides was higher than in the general population—no doubt partly because a man's illness prevented him from working and partly because dismissal from work had helped to provoke gloom or despair

For convincing reasons, Sainsbury concluded that a lonely way of life and some social disorganization favour the occurrence of suicide Taken by itself it is not a startling conclusion, but it must be seen alongside the negative findings—for example that there was no intrinsic connection between overcrowding and suicide, though it was previously thought that there is, and that drift of predisposed persons into the areas that had high suicide rates did not account for these rates

An ecological study has serious limitations It presupposes social homogeneity in the ecological area (Clausen and Kohn, 1954) It cannot penetrate into the details of the nexus between social variables and mental illness rate It can only in exceptionally favourable circumstances enable us to determine which

social variables in an area are responsible for that area s rate of mental disorder It can, however, indicate where a close search into a presumptive causal relationship is well warranted and where it has little prospect of reward It does this best when it deals with a clinical anomaly that can be fairly well defined (like suicide), when it has areas whose relevant social char acteristics are distinctive and readily ascertainable and most important of all, where there is a well founded hypothesis to be tested

In Sainsbury s study the main hypothesis and the method derived from Emile Durkheim (Durkheim, 1951) Durkheim had divided the social causes of suicide into three classes those which promote excessive individualism and isolation secondly, those which exercise compelling authority over the individual, driving him into conforming or altruistic acts such as suttee or giving one s life for another and thirdly, those which bespeak social disorganization—anomic forces as Durkheim called them, which leave the individual unsupported by the social controls and assurances that should keep him steady and secure This threefold classification by Durkheim has commended itself to many subsequent workers, though they have mostly paid more attention to the psychological aspects of the matter than he did

Sainsbury found social isolation to be the predominant social factor in the populations he studied A similar conclusion could be drawn from a more exotic study made in Singapore (Murphy, 1954) The population of Singapore falls into five main ethnic groups—Chinese Indian European, Eurasian and Malay, each with a different suicide rate The Chinese, who make up three quarters of the total population, have a high suicide rate (21 per 100 000) and so have the Indians whereas the Malays (like Muslims everywhere) have a low rate—only 1 7 suicides per 100 000 of population Within the Chinese community there are wide variations in suicide rate (from 12 to 24 in men and from 6 to 16 per 100 000 in women) and these variations are associated with tribal factors the Chinese are immigrants from five different regions of China, and their place of origin can be recognized from the dialect they speak The smaller and more economically insecure the dialectal sub

group, the higher the suicidal rate tended to be Similarly the suicide rate was higher in areas where there was most admixture of the dialectal sub groups, betokening the insecurity of a transplanted minority Though isolation was evidently a factor in the causation of suicide, it had less effect (according to Murphy, who made this investigation) upon suicide itself than upon the frequency of manic depressive psychosis (of which, of course suicide may be an outcome) Murphy s study in Singapore, to which I have just referred, is a striking example of the application of ecological methods to the study of an ethnically mixed population

Cultural Influence

This brings us to the anthropological methods and issues that are relevant to the social causes of mental disorder (Hsu, 1951, Opler, 1956) The anthropologists have forced us to re examine our notions of what is normal and healthy They have demonstrated that behaviour which in our culture would indicate mental disorder is in other cultures acceptable and healthy The point has been so often made and is so obvious that I need not labour it It has reinforced the psychiatric maxim that aberrant behaviour can never be assessed in a social vacuum, but must always be related to its background in time and space *the 'here and now standpoint is, by itself, misleading* Conversely it is by no means easy to decide whether a particular person, living in a culture alien to him or in a culture alien to the examining psychiatrist is mentally ill or not The anthropologist is primarily concerned with groups of normal individuals who fulfil culturally approved roles whereas the psychiatrist is mostly preoccupied with individuals who are somehow at odds with their cultural environment

The great difficulties which beset the collection and comparison of statistics about prevalence of psychiatric disorders in European, North American and other Western societies are immensely increased when we wish to compare the frequency of mental disorder, and its particular varieties, in Asiatic or African communities with those in our own countries Attempts have of course been made to do this even on the evidence

available from non literate, primitive communities where such things as adequate mental hospital provision and censuses are unknown But the data so presented are untrustworthy, except in the few instances where competent investigators have carried out the study

An example of the latter is the Formosan inquiry into incidence made by Dr Tsung Yi Lin in 1946–8 (Lin 1953) Conditions were rather favourable The society is homogeneous, except for a few large cities and the mountainous districts where the aborigines live The population has been static for several generations, and falls into two slightly different ethnic groups, both from South China The investigators were competent and were themselves Formosans Yet even so they concluded that apart from the major psychoses and epilepsy they could not make any reliable comparison between their estimates of the incidence of neurosis psychopathic personality, or alcoholism and those available from other communities For psychoses the rate they found was lower (3 8 per 1 000) than investigators have reported from other countries But the chief interest of this study lay in its implications for the mental health services The investigators concluded that the family plays so large a part in influencing behaviour in this Chinese culture that the family pattern should be preserved when mental illness brings the patient to hospital Consequently the wife or mother of a patient became a nursing aide so that family care was extended to the ward earlier discharge was made possible and the responsible relative learnt how to deal with the patient after he returned home So we find an ecological survey of mental illness in Formosa leading to an arrangement very similar to that arrived at in another way for children admitted to English hospitals

The diverse forms of mental disorder in different cultures have been of interest to many psychiatrists including Kraepelin (1904) Some of the manifestations—like running amok or Latah among the Malays—have become very widely known These varieties of mental illness strange to us, could be taken to indicate either inherited differences between Malays and ourselves, or cultural differences That they rest on specific

group, the higher the suicidal rate tended to be Similarly the suicide rate was higher in areas where there was most admix ture of the dialectal sub groups, betokening the insecurity of a transplanted minority Though isolation was evidently a factor in the causation of suicide, it had less effect (according to Murphy, who made this investigation) upon suicide itself than upon the frequency of manic depressive psychosis (of which, of course, suicide may be an outcome) Murphy's study in Singapore to which I have just referred, is a striking example of the application of ecological methods to the study of an ethnically mixed population

Cultural Influence

This brings us to the anthropological methods and issues that are relevant to the social causes of mental disorder (Hsu 1951, Opler, 1956) The anthropologists have forced us to re examine our notions of what is normal and healthy They have demon strated that behaviour which in our culture would indicate mental disorder is in other cultures acceptable and healthy The point has been so often made and is so obvious that I need not labour it It has reinforced the psychiatric maxim that aber rant behaviour can never be assessed in a social vacuum, but must always be related to its background in time and space the here and now standpoint is, by itself, misleading Con versely it is by no means easy to decide whether a particular person, living in a culture alien to him or in a culture alien to the examining psychiatrist is mentally ill or not The anthro pologist is primarily concerned with groups of normal indivi duals who fulfil culturally approved roles whereas the psychia trist is mostly preoccupied with individuals who are somehow at odds with their cultural environment

The great difficulties which beset the collection and compari son of statistics about prevalence of psychiatric disorders in European, North American and other Western' societies are immensely increased when we wish to compare the frequency of mental disorder, and its particular varieties, in Asiastic or African communities with those in our own countries Attempts have, of course, been made to do this even on the evidence

of the psychotic reactions of individuals with different racial and cultural backgrounds, differences which make it impossible to fit them into the accepted nosological framework To examine this conclusion properly, we should have to consider the basis on which we differentiate our accepted nosological classes, but apart from this fundamental question, the illustrative cases and comments which Tooth provides make his conclusion questionable It is, at all events, clear from what he tells us in his Report that social influences play a notable part in determining the course and form of mental illness Thus apropos of atypical psychoses in which there were delusions hallucinations, and shallow affect, he writes Most of the individuals came from the Northern Territory and those from the South were found in small villages where their lives, both before and during their illnesses, were relatively simple and sheltered In only a few instances had any form of restraint been used and then only by the family within the family compound As soon as the acute phase had passed, the patient was encouraged to carry on with his usual activities and, by virtue of his peculiarities he acquired a certain standing in the community Thus without losing the background of family security, he was kept occupied in familiar surroundings, a situation which mental hospitals all over the world are seeking to achieve by artificial means Contrasting these people living in their homes with those who have strayed to the large towns and those who have been confined to the asylum one is struck by the dilapidation of the former and the restlessness and florid symptomatology of the latter It is suggested that these endogenous psychotics left to fend for themselves outside the family would have developed into the dilapidated eccentrics of the market place but if transported into an alien and restricted environment would have produced the symptomatology of one of the familiar varieties of schizophrenia In short the form of psychosis had been determined by the treatment

The extent to which the form of a psychosis can be affected by the cultural and climatic conditions of a people is well illustrated by the Windigo psychosis of the Chippewa Ojibwa and Cree Indians (Landes 1938 Linton, 1956) These people suffer

inherited tendencies is very unlikely If, as we may safely assume, they rest on cultural influences, it is illuminating to see which of our forms of mental illness they correspond to, and how the particular culture has created their characteristic features

There is a wide range of these unusual conditions Witigo among Canadian Indian tribes, Koro in South eastern Asia and Borneo, Candomble possession in Bahia (Stainbrook, 1952), and other oddly named forms of very queer behaviour have been described in many parts of the world The best known, and the prototype of many of them, is Latah it is a reaction provoked by sudden fright and signalized by automatic obedience to commands, imitation of words and gestures, and obscene ejaculations It occurs in naive, poorly endowed persons, and among the Malayans is not regarded as much more than an eccentricity, unless it is severe It occurs not only in Malaya, but also in Java, the Philippine Islands, Borneo, Burma, Siam, India, Siberia and Lapland, Japan and Africa (Yap, 1952) Always it affects the most neglected or backward people, such as the Ainu in Japan, and many of those exhibiting it show evidence of mild dementia or chronic anxiety It has often been cited as an example of a special disease, restricted to particular ethnic groups This view can no longer be sustained It is closely parallel to conduct observed among some American settlers of European origin, and it is similar to the behaviour of superstitious and illiterate people in many times and countries when under the influence of religious excitement The more careful the clinical study of the condition, the more does it fit into the categories of conventional psychiatry, the more careful the anthropological study the more plainly does it reveal cultural influences working upon suggestible often defective, psychopathic subjects

There is, however, a risk in putting too much stress on the cultural aspect of strange forms of mental disorder A congeries of atypical mental conditions seen on the Gold Coast turned out to be manifestations of trypanosomiasis (Tooth, 1950) But where physical disease had been excluded there were still atypical psychoses, indeed Tooth says One is therefore forced to the conclusion that there are real differences in the quality

generalize and apply what we learn from studying them But the contrast is not so glaring as it seems Anthropologists have shown that it is false to picture primitive societies as relatively uniform aggregations of custom ridden savages and we know well enough how many irrational beliefs and how much irrational conduct can be found in Western societies Moreover, culture (in the sense in which anthropologists use the term) covers not merely the social organization, the way of life of a society but also its symbolism, its language, art, gestures and other modes of expression and communication The core of a people s culture lies in ideas, values and habits that have been historically selected over a long period It moulds the individual, brings him into relation with others and of course may itself be modified by what he and other individuals originate or contribute to it Every cultural system is complex and cultural similarities turn up in the most unlikely places The deficiencies in psychiatric studies of primitive cultures lie not in the place or method of study but in the imperfect knowledge we have of the details of the psychiatric disorders and of the social structure and beliefs in the societies in question It is hard enough to find out all we should know about mental illness and its background here, where we are at home and have access to so much information when we come to Latah or Arctic hysteria or Koro or Piblokto among the Eskimos of Greenland we have only approximate information mostly collected by people with insufficient expert knowledge of what is required

A special cultural feature to which much attention has been paid is the child rearing practices of a given society It is of course a commonplace that childhood is the most plastic and impressionable period of life Psychoanalytic views put particular stress on certain phases and experiences of childhood as determinants of subsequent personality and of mental abnormalities A great deal has in consequence been written about the effects of various sorts of infant training and of deprivation of mother love (Bowlby 1951 Erikson, 1951) Critical examination of this literature (Orlansky 1949 Lindesmith and Strauss 1950 Hsu 1951 Pinneau 1955 Opler 1955) shows that it is loosely stated and much of the evidence is inadequate to bear

terrible hardships in the severe winters of north east Canada The scarcity of game obliges each family to live by itself exposed to the risk of starvation, cannibalism sometimes occurs They have myths about a monster, living as an ice skeleton during winter and dying in the spring, who devours human beings They believe also that human beings may be led by witchcraft, to develop similar cannibal desires and to have their heart turn to ice In fact some members of the tribe do become profoundly depressed and excessively anxious about starvation Their perceptions then become disturbed and they see members of their family as plump, succulent, inviting beavers Some of those affected have insight into their condition and beg that they should be killed before they give way to their cannibal urge, others actually kill and eat members of their family and eventually other people if they are not caught in time Normally, if I may use the word in this context, the tribe seize the patient as soon as he has eaten anyone, and kill him—burning his corpse in order to destroy the hidden ice which they believe to be so potent

The broad connection here between the physical privations and risks and the form of insanity in these tribes is evident, but as Linton points out, it would be false to attribute this psychotic behaviour entirely to the stresses of the environment There are many other places where tribes face as grave a risk of cold and starvation as these Canadian Indians, but they have no such cannibalistic psychoses Thus, where is hunger and ice are ever present threats also in other areas this particular psychotic phrasing of the manifestations of ultimate despair is a distinctively localized and cultural phenomenon closely correlated with many aspects of Chippewa mythology, shamanism, social organization economy and habitat

There are many reports about psychotic behaviour in primitive societies which similarly show how the culture of a people colours and shapes the clinical picture of insanity It may be objected however, that such findings have little or no relevance to our problems these crude people with their fantastic superstitions and simple way of life are too remote from our highly organized, rational, individualistic society to permit us to

scientists Unfortunately this is not at all the case With the exception of a few papers there is as yet practically no research literature on the impact of mental illness on the family (Clausen and Yarrow, 1955) Nevertheless there is no dearth of opinions based on clinical experience and current assumptions about family relations It is in the light of these unconfirmed opinions that the clinician usually acts when he attempts to rehabilitate and resettle, say, a schizophrenic who had been for years in a mental hospital and who still has residual symptoms (Carstairs O Connor and Rawnsley 1956) The empirical procedure is often successful But there is an obvious gulf between these more or less intuitive efforts and the application of assured knowledge about the social meaning, social effects and social possibilities of treatment in mental illness

The impact of mental illness is of course, not only on the family, but on wider circles—on the fellow workers and other associates of the patient, on those who entrust themselves to him (in the way that patients put themselves in the hands of a doctor, soldiers of a general or passengers of an engine driver), and mental illness has its impact finally on society at large—a paranoid assassin a psychopathic dictator a melancholic religious leader can exert a far reaching effect even on generations unborn and mental illness as an economic and psychological burden, has its cumulative effect on society

Without ranging too far in place and time it is useful to examine the extra familial aspects of social adjustment to mental disorder, and also to mental defect It was in fact work on mental defect by the Social Psychiatry Research Unit (Medical Research Council 1953, O Connor and Tizard, 1956) that paved the way for the studies we are now making in the social relations of schizophrenia It is impossible to develop a sound programme for the social and occupational betterment of high grade defectives or imbeciles unless one has first examined and as far as possible measured, the abilities of the affected population and then organized a training procedure which takes into account these estimates of ability to perform tasks and to learn new ones and which prepares the defectives under experimental conditions to respond adequately to social

the conclusions drawn The main inference that it is beneficial for a child to have been kindly and affectionately brought up needs no reinforcement since it conforms to universal experience But when the social implications come to be considered, there is a dearth of well designed, properly evaluated studies On the whole this has been, in proportion to the amount written about it, so far an unrewarding field of research in social psychiatry

The Family

Family relationships in general, however, are a ground of common concern where the professional interests of psychiatrist, sociologist and anthropologist meet For the sociologist the family is a nuclear unit, for the anthropologist bonds of kinship are of the first importance, and for the psychiatrist the family is the matrix within which the individual is moulded and developed, the area where his strongest emotional ties are formed, the background against which much of his most intense personal life is enacted There is therefore need to study the family, not only from the psychoanalytical and psychological standpoint but also to discover how mental illness impinges upon it, and what effects this sort of incapacity has on the family structure

The problem is obvious enough when neurotic disorders like hysteria or hypochondriasis wear down the patience and affection of other members of the patient s family, or when psychopathic personality alcoholism or sexual perversions play havoc with family life Schizophrenia, which may entail many years of continuous mental hospital care leads to far reaching changes in the family, which are important in assessing prognosis and especially in regard to the chances of social rehabilitation

The impact of mental illness especially schizophrenia, upon the family is now being studied by joint groups of psychiatrists and social psychologists in Washington and also in London But as one of the investigators has recently put it One might anticipate that so substantial a practical problem, with so many aspects of significance for understanding the dynamics of human relationships, would have been thoroughly studied by social

These developments are signs of the interplay between social forces and psychiatry They are not all equally sound and welcome Public expectations are sometimes, unfortunately, fed with information that is incomplete or ill founded It will be salutary when this is no longer so More knowledge of the social as well as of the somatic and psychological processes that operate in psychiatry, will prevent unreal hopes being fostered and permit more effective measures of prevention and amelioration than we yet have

REFERENCES

BOWLBY J (1951) *Maternal Care and Mental Health* World Health Organization Geneva

BRUGGER, C (1931) *Z ges Neurol Psychiat* **133** 352

BRUGGER C (1938) *Ibidem* **160** 189

BUCKLE H T (1857) *History of Civilization in England* Longmans London

CARSTAIRS G M O CONNOR N and RAWNSLEY K (1956) *Brit J prev soc Med* **10** 136

CLAUSEN J A and KOHN M L (1954) *Amer J Sociol* **60** 140

CLAUSEN J A and YARROW M R (1955) *J soc Issues* **11** 3

DURKHEIM E (1951) *Suicide A Study in Sociology* (Tr J A Spaulding and G Simpson) Free Press Glencoe Ill

EATON J W and WEIL R J (1955) *Culture and Mental Disorders A Comparative Study of the Hutterites and Other Populations* Free Press Glencoe Ill

ERIKSON E (1951) *Childhood and Society* Imago London

ESQUIROL E (1838) *Des Maladies Mentales* Brussels

FARIS R E L and DUNHAM H WARREN (1939) *Mental Disorders in Urban Areas* University Press Chicago

GERARD D L and HUSTON L (1953) *Psychiat Quart* **27** 90

GOLDHAMER H and MARSHALL A W (1953) *Psychosis and Civilization* Free Press Glencoe Ill

HOLLINGSHEAD A B and REDLICH F C (1954) *Amer sociol Rev* **19** 302

HSU F L K (1951) *Southwestern J Anthropol* **8** 227

KRAEPELIN E (1904) *Z blatt Nervenheilk* **15** 433

LANDES R (1938) *J ab soc Psychol* **33** 14

LEWIS A (1953) *Brit. J Sociol* **4** 109

LIN T (1953) *Psychiatry* **16** 313

LINDESMITH A R and STRAUSS A L (1950) *Amer sociol Rev* **15** 587

LINTON R (1956) *Culture and Mental Disorder* Thomas Springfield

MARSHALL H (1953) *Mental Health Statistics* Dominion Bureau of Statistics Ottawa

requirements and to carry social responsibility The preparation has included group psychological treatment, and a graduated series of rewards and responsibilities, also close examination of the situation in the patients' families and at their prospective residences and places of work Psychological, clinical and social studies have thus gone hand in hand The result has been impressive At the hospital where the work was carried on many defective patients were resettled in the general community and, in financial terms, the change was from £8,131, earned by the patients on daily licence, in 1948 to £30,000 in 1951 But the result cannot be measured by earnings, but by the well being of those defectives who return to lead a fairly normal and happy existence in society Though satisfying, the fruits of this research point to the necessity for further socio psychological inquiry directed not only at those who have been certified and admitted to an institution but also at the numerous patients who have all along been cared for at home

Social Aspects of Treatment

The social aspects of psychiatric treatment are manifold The attitude a society has towards mental disturbance and its treatment is itself an important, far from static, social force, determining how far existing facilities are availed of and how far different or more extensive facilities are demanded Swayed by the prevailing social attitudes, the patient or his relatives often nowadays put pressure on the psychiatrist to provide electrical convulsant therapy, or tranquillizing drugs or leucotomy, or psychoanalysis, conversely, the social climate may make them refuse a particular treatment Leucotomy is legally forbidden in at least one large country, and frowned on in others, for reasons that have little or nothing to do with its therapeutic efficacy

There are many current developments the growth of day hospitals and of interest in mental hospitals as therapeutic communities, the increasing public concern for the corrective treatment of juvenile and other offenders, the public eagerness to know about psychopathology and therapy and the desire of alarmed parents to be told how they can bring up their children without doing them some irreparable psychological harm

These developments are signs of the interplay between social forces and psychiatry They are not all equally sound and wel come Public expectations are sometimes unfortunately fed with information that is incomplete or ill founded It will be salutary when this is no longer so More knowledge of the social, as well as of the somatic and psychological processes that operate in psychiatry, will prevent unreal hopes being fostered *and permit more effective measures of prevention and ameliora* tion than we yet have

REFERENCES

BOWLBY J (1951) *Maternal Care and Mental Health* World Health Organi zation Geneva

BRUGGER C (1931) *Z ges Neurol Psychiat* **133** 352

BRUGGER C (1938) *Ibidem* **160** 189

BUCKLE H T (1857) *History of Civilization in England* Longmans London

CARSTAIRS G M O CONNOR N and RAWNSLEY K (1956) *Brit J prev soc Med* **10** 136

CLAUSEN J A and KOHN M L (1954) *Amer J Sociol* **60** 140

CLAUSEN J A and YARROW M R (1955) *J soc Issues* **11** 3

DURKHEIM E (1951) *Suicide A Study in Sociology* (Tr J A Spaulding and G Simpson) Free Press Glencoe Ill

EATON J W and WEIL R J (1955) *Culture and Mental Disorders A Comparative Study of the Hutterites and Other Populations* Free Press Glencoe Ill

ERIKSON E (1951) *Childhood and Society* Imago London

ESQUIROL E (1838) *Des Maladies Mentales* Brussels

FARIS R E L and DUNHAM H WARREN (1939) *Mental Disorders in Urban Areas* University Press Chicago

GERARD D L and HUSTON L (1953) *Psychiat Quart* **27** 90

GOLDHAMER H and MARSHALL A W (1953) *Psychosis and Civilization* Free Press Glencoe Ill

HOLLINGSHEAD A B and REDLICH F C (1954) *Amer sociol Rev* **19** 302

HSU F L K (1951) *Southwestern J Anthropol* **8** 227

KRAEPELIN E (1904) *Z blatt Nervenheilk* **15** 433

LANDES R (1938) *J ab soc Psychol* **33** 14

LEWIS A (1953) *Brit J Sociol* **4** 109

LIN T (1953) *Psychiatry* **16** 313

LINDESMITH A R and STRAUSS A L (1950) *Amer sociol Rev* **15** 587

LINTON R (1956) *Culture and Mental Disorder* Thomas Springfield

MARSHALL H (1953) *Mental Health Statistics* Dominion Bureau of Statis tics Ottawa

MEDICAL RESEARCH COUNCIL (1953) Report for the Year 1951– H M S O London
MOORE T V (1936) *Ecclesiastical Rev* **95** 486 and 601
MURPHY H B M (1954) *Med J Malaya* **9** 1
NORRIS V (1956) *J ment Sci* **102** 467
O CONNOR N and TIZARD J (1956) *The Social Problem of Mental Deficiency* Pergamon Press London
ØDEGAARD Ø (1932) *Emigration and Insanity* Levin and Munksgaard Copenhagen
ØDEGAARD Ø (1946) *J ment Sci* **92** 35
OPLER M K (1955) *Soc Problems* **3** 12
OPLER M K (1956) *Culture Psychiatry and Human Values* Thomas Springfield
ORLANSKY H (1949) *Psychol Bull* **46** 1
PINNEAU S R (1955) *Psychol Bull* **52** 4 9
REGISTRAR GENERAL (1955) *Statistical Review of England & Wales Mental Health 1950–1951* H M S O London
SAINSBURY P (1955) *Suicide in London An Ecological Study* Maudsley Monograph Series Chapman and Hall London
SIMMONS L W and WOLFF H G (1954) *Social Science and Medicine* Russell Sage Foundation New York
STAINBROOK E (195?) *Amer J Psychiat* **109** 330
STARKEY M L (1952) *The Devil in Massachusetts* Hale London
STENGEL E (1958) *Attempted Suicide* Maudsley Monograph Series Chapman and Hall London
TOOTH G (1950) *Studies in Mental Illness in the Gold Coast* H M S O London
YAP P M (1952) *J ment Sci* **98** 515

IX

Causes of Pain

C A KEELE

My purpose in this lecture is to discuss the modes of stimulation of pain nerve endings and the peripheral processes responsible for pain production in disease states

I begin by accepting that pain is a specific mode of sensibility distinct from other sensory modalities and I assume that it is mediated by free nerve endings and a separate nerve pathway The pain nerve endings may be compared with nerve receptors of other kinds In the skin itself there are receptors particularly sensitive to temperature, touch and pressure, though Weddell s recent work (see Weddell Palmer and Pallie 1955) questions the existence of specific histological structures for these modali ties and in other parts of the body there are for example the baroreceptors of the carotid sinus, the chemoreceptors of the carotid body and the osmoreceptors and thermoreceptors in the hypothalamus

Pain nerve receptors differ from those just mentioned in being more versatile, in responding to a wide range of physical and chemical stimuli To adapt a phrase from another field they have a broad spectrum of sensibility It is not known whether the pain nerve endings which respond to physical stimuli are the same as those which respond to chemical stimuli

Pain receptors have a high threshold compared with other receptors and this is really most appropriate for life would be quite intolerable if pain receptors were as sensitive as other receptors In that case the lightest touch the smallest changes in temperature, and the slightest deviations from normal in the

CO tension, pH and osmolarity of the body fluids would summate to produce constant and unbearable pain

Actually, as Walshe has pointed out (1948), pain receptors are on the whole well fitted for their task, which would appear to be to signal impending or incipient damage to the body Cutaneous pain leads to reflex and voluntary evasive action which helps to reduce the damage from the nocuous stimulus to a minimum Sir Thomas Lewis (1942) and Hardy, Wolff and Goodell (1952) have emphasized that when pain is felt some tissue damage has already occurred, and that tissue damage may sometimes be produced without pain being felt at all This is particularly the case with regard to tumours which may grow for some time without being painful

Pain due to Pressure or Tension

Apart from the pain produced by excessive pressure or tension applied to the surface of the body e g a blow on the shin or the tweaking of a hair, there are also many examples of visceral pain arising in this way The pain of childbirth, which is associated with powerful uterine contractions, the colics of the alimentary, biliary and urinary tracts are all due to the tension exerted on pain nerve endings and there is much evidence which shows that stretching of the walls of blood vessels by excessive dilatation or by traction can also produce pain

Investigations on the mechanisms involved in the production of headache, by Pickering (1939) and by Wolff (1948), have shown the great importance of afferent vascular innervation in this type of pain Pickering and Hess (1933) showed that the throbbing headache which follows injection of histamine is due to dilatation and distension of intracranial arteries and can be abolished by overcoming the distension by raising the intracranial pressure Scott and Warin (1946–8) have found that the headache associated with fever is of similar origin and so probably is the well known hangover headache of the morning after

Changes in *intracranial pressure* may produce headache by stretching the walls of intracranial arteries or large venous sinuses Wolff found that experimental increase of intracranial

pressure in normal subjects did not produce headache whereas lowering of intracranial pressure by removal of 20 ml of cerebrospinal fluid regularly did so, particularly when the subject assumed the erect position This type of headache was enhanced by compressing the jugular veins and relieved by restoration of cerebrospinal fluid volume Headache following *lumbar puncture*, which is probably due to leakage of cerebrospinal fluid clearly resembles that which develops after experimental drainage

Wolff observed that in patients with cerebral tumour headache occurred almost as frequently with normal as with increased intracranial pressure and in some instances could be induced by lowering this pressure He concludes that the headaches produced by drainage or by cerebral tumour are not due to changes in intracranial pressure *per se* but to dilatation of and traction on pain sensitive intracranial vascular structures though tumours may of course also press directly on cranial nerves

Migraine Wolff has shown that in migraine the development of the characteristic unilateral headache is associated with distension and increased amplitude of pulsation of the branches of the external carotid arteries on the affected side and that injection of ergotamine reduces the amplitude of arterial pulsation at the same time as the intensity of headache decreases Ostfeld and Wolff (1955) have recently reported that infusion of noradrenaline also relieves the headache of migraine

Thus stretching and increase of tension in hollow viscera and the walls of large blood vessels can produce pain when the threshold for pain in the involved nerve endings is normal When there is inflammation the pain threshold is lowered and degrees of tension which would normally not produce any unpleasant feeling can be exquisitely painful as in the throbbing pain due to arterial pulsation in an abscess I shall refer later on to factors which enhance the sensitivity of pain nerve endings in such cases

Now although we commonly think of physical stimuli as acting directly on pain nerve terminals it is quite possible that some such stimuli act via chemical processes which in turn

excite the nerve endings' For example, Hardy *et al* (1950) have pointed out that application of heat to the skin only produces pain when the temperature reaches 45 C, and that at this temperature denaturation of proteins begins to develop However, the speed of onset of such pain would make the study of associated chemical changes very difficult and the possibility of chemical mediation of such a stimulus must remain speculative for the time being

The pain associated with compression of nerves for example by a tumour, a prolapsed intervertebral disc or an aneurysm, raises interesting problems If compression of a nerve is great enough conduction will finally be blocked and there will be loss of all sensation in the innervated region Pain can occur only with lesser degrees of compression or when compression is intermittent The importance of the blood supply to nerves has been emphasized by many workers in recent years who have shown that ischaemia of nerves increases their irritability, producing paraesthesiae and hyperalgesia Thus compression probably produces pain by reducing the blood flow through the nerve, the consequent hyperalgesia so increasing the sensitivity that normally subthreshold stimuli in the affected area become painful The physical stimulus of pressure thus produces pain by an indirect process, by inducing metabolic changes in the ischaemic fibres but the nature of these changes is unknown

And while on the subject of ischaemia of nerves I should like to mention the pain of *causalgia* which follows injury to nerve trunks Doupe, Cullen and Chance (1944) have made the very interesting suggestion that in this condition chronic ischaemia of the affected nerve leads to a breakdown of the insulation between nerve fibres so that efferent impulses passing down the postganglionic sympathetic nerve fibres can traverse artificial synapses in the damaged region and set up afferent impulses in neighbouring pain nerve fibres This view is supported by the recorded beneficial effects of sympathectomy in this condition Richards (1951) has suggested that the pain of ischaemic neuritis associated with obstructive arterial disease is of similar origin

Chemical Factors in the Production of Pain

The production of pain by chemical factors has been little studied experimentally because it is difficult to make repeated applications of the substances under investigation, and to obtain accurate measurements of the results We know of course that surface application of strong irritants such as chloroform acids and alkalis produces severe and painful damage in the skin and mucous membranes and that injected solutions of certain drugs, such as organic arsenicals and mercurials and of hypo or hypertonic solutions will produce pain in both skin and deeper structures Other examples of *extrinsic* pain producing materials are nettle stings and certain venoms, such as wasp venom, which I shall mention later

It is, however more important to consider the possibility of pain production by *intrinsic* chemicals substances derived from living body fluids and tissues Consider for example, the pain of *peptic ulcer* It is natural to suspect that pain in this condition is due to the action of the acid gastric juice on the ulcerated surface, and the relief of this pain by administration of antacids supports this view Moreover Palmer (1926) and Bonney and Pickering (1946–8) have reported that administration of hydrochloric acid at pH of 1–2 produces pain in many patients with ulcer, and that when spontaneous ulcer pain occurs the pH of the gastric contents is markedly acid I shall have further comments to make on this later on

To take another illustration—the pain and tenderness in *ischaemic muscle* MacWilliam and Webster (1923) suggested that the pain which develops when muscles are worked under ischaemic conditions might be due to metabolic products of activity Lewis Pickering and Rothschild (1931) developed this view and postulated a factor P as the stimulus to pain Factor P is presumed to arise from metabolic activities in muscle fibres and to diffuse thence to the extracellular space where stimulation of pain nerve endings takes place The steady increase of pain during ischaemic muscular work is attributed to accumulation of factor P in the tissue spaces and the very rapid relief which follows restoration of the circulation to the speedy dispersal of this sub

stance Lewis did not identify factor P and suggestions by others that it is lactic acid or potassium have never been confirmed

The chief interest in this type of pain lies in the explanation which it offers for the pain which accompanies activity in skeletal and cardiac muscle in the presence of obstructive arterial disease The pain of intermittent claudication in the limbs, and of angina pectoris, are both typically provoked by exertion and relieved by rest With sudden complete occlusion of an artery, as in coronary thrombosis, the onset of pain is similarly produced but it cannot of course be relieved by rest

Inflammation is another state in which it has been postulated that pain is provoked by chemical agents Dolor was one of the cardinal features of inflammation mentioned by Celsus, the others being rubor, calor and tumor Whilst much work has been done on the factors which may account for the vasodilatation and increased permeability of the inflammatory response, much less attention has been paid to the processes by which pain arises in this state, though it is commonly said that the pain is due to increased tension in the inflamed tissues Now whilst there may certainly be an increased tension in an inflamed area this factor cannot directly provide more than a partial explanation since the outstanding feature of an inflammatory process is the marked *hyperalgesia* which it induces In inflamed tissues the pain threshold is much reduced and stimuli which are not painful in normal tissues may be very painful in the affected region The hyperalgesic state may be such that even arterial pulsation can produce a severe throbbing pain with each heart beat

Sometimes an inflamed area is spontaneously painful, that is pain is present without application of an external stimulus or when the inflamed part is completely at rest Lewis (1942) suggested that in such cases the chemical agents which produce hyperalgesia have reached a sufficiently high concentration to stimulate the pain nerve endings on their own

One of the examples which Lewis gave in support of the idea of chemically induced pain was the sequence of events following a heat burn I would like to illustrate this in relation to some of our own investigations

Figure 1 shows how such pain may arise after experimental burns to the skin The figure shows graphic records in two subjects of the pain produced by application of heat at 65 C and 70° C, in each case for 3 seconds The initial high peak of pain

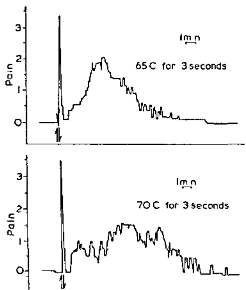


FIG 1 Records of pain following experimental burns The burns were induced in the skin of the forearm by application of a heated brass rod 1 cm in diameter at 65 C in one subject and 70 C in the other the duration of application being 3 seconds in both cases Pain was recorded as described by Armstrong *et al* (1953) Ordinate = Pain Scale where 1 = slight pain 2 = moderate pain 3 = severe pain and 4 = very severe pain Note the peak of pain during application of heat followed by latent period of 27 and 21 seconds respectively when no pain was felt and then a slowly waxing and waning of what Lewis called the delayed pain of a burn

coincided with the period of heating after which the pain rapidly subsided to or almost to zero Then after 21–27 seconds pain returned and rose more slowly to a moderate intensity and then decayed again during a period of about 10 minutes after which the area was hyperalgesic but not spontaneously painful It is the delayed pain which Lewis ascribed to chemical factors and I shall later consider what substances might be involved

The subsequent hyperalgesia resembles that produced by other forms of injury such as occurs after excessive cooling (−15 C), exposure to ultraviolet light, scratching and infective and allergic inflammatory processes

Lewis showed that when there was cutaneous hyperalgesia following injury spontaneous pain could be aroused either by warming the skin to 30–35° C or by occluding the circulation to the affected part He suggested that in both these cases the spontaneous pain was produced by an increased local concentration of metabolites, the raised temperature increasing their production and the ischaemia preventing their removal However, as pointed out by Nathan (1953) occlusion might act by increasing the sensitivity of the pain nerve fibres, so that pain could occur without any increase in concentration of pain producing metabolites

So far there have been no suggestions as to what these metabolites are Lewis did not commit himself except to say that he did not think histamine was concerned He found that aqueous extracts of skin caused pain on intradermal injection but did not identify the active principle

I will now deal with work which we have been doing at the Middlesex Hospital Medical School My colleagues in this work are Dr Desiree Armstrong, Dr J B Jepson and Dr J W Stewart, and together we have been investigating the possible role of intrinsic chemical factors in the production of cutaneous pain with particular reference to substances which develop in blood

In conjunction with Drs Dry and Markham (Armstrong, Dry Keele and Markham, 1953) we first studied various methods for producing pain by chemical agents Lewis had administered histamine and other substances by *pricking through a drop of solution* placed on the skin, but we soon found that this *mode of application was unsatisfactory* since pain was not produced by substances which were very active when given in other ways We also tried *intradermal injection* which was certainly effective with most substances which we tested However, we sometimes observed irregular responses such that isotonic saline might occasionally be painful and normally painful solutions ineffective Of course with sufficiently numerous injections

such irregularities could be taken account of but our subjects (I was one of them!) did not seem to enjoy receiving the large number of injections needed to give statistically significant results Further, it was obviously dangerous to inject inflammatory exudates which might be contaminated with bacteria or foreign materials to which the subjects might become sensitized So we sought and found, what was for our purposes a better method, in which the substances to be tested are applied to the exposed base of a blister produced by cantharidin

Blister Area A 0 3 per cent cantharidin plaster about 1 cm in diameter is applied to the skin of the forearm for 4–6 hours in the evening By the next morning a blister has usually developed The fluid (about 0 2 ml) is aspirated, the raised epidermis is removed and the base of the blister is then available for testing Dr J Boss has shown that all but the basal epidermal cells are removed and that the cells of this layer tend to be separated from each other though still attached to the basement membrane The nerve fibres which end in the epidermis will be ruptured during blister formation but deeper nerve terminals will remain intact

The area is bathed in Locke s solution which does not produce pain This solution is removed at intervals (usually 10 minutes) to allow application of the test materials We have found this technique to be very reliable and sensitive and the effect of a given solution is reproducible during an experimental session lasting 1½–2 hours The subject of course does not know the nature of the applied materials and on some occasions the experimenter has also been kept in ignorance The results are recorded graphically in the manner already described

We are aware that the blister base is not a normal tissue and that it might be hypersensitive to certain agents but since pain often occurs in already damaged tissues the conditions of our tests are reasonably relevant The constancy of response to a given concentration of a standard agent is of great value in assessing the rise and fall in activity of labile pain producing fluids such as inflammatory exudates and plasma Solutions need not be sterilized as they are rapidly removed but we have naturally avoided heavily contaminated materials

We have obtained much useful information from the blister area which has most of the advantages of an isolated organ preparation combined with the very unusual property of remaining attached to the rest of the body

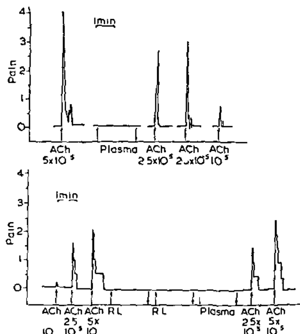


FIG 2 Production of pain by acetylcholine in two subjects Solutions of ACh were applied to the exposed base of a cantharidin blister Ordinate = subjective pain Scale as described in Fig 1 RL = Ringer Locke solution Numbers below ACh show amounts in g/ml of fluid (Armstrong *et al* 1953)

I will now describe the effects of various pain producing agents as tested on the blister area

Acetylcholine (ACh) When different amounts of ACh (in Locke s solution) are applied to the blister area pain is produced with concentrations from 10^{-5}g /ml upwards The pain is commonly immediate in onset smarting or stinging in character, rapidly reaches a peak and then subsides to zero within 30–60 seconds The effects of different doses are graded (Figure 2)

Potassium Chloride Solutions of potassium chloride (0 1–1 per cent) produce pain rather similar to that seen with ACh but the effect comes on a little more slowly and lasts rather longer

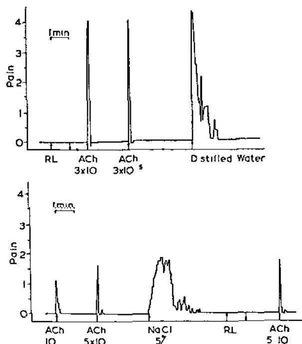


FIG 3 Production of pain in blister base by distilled water and 5 per cent NaCl solution Ordinate = Pain Scale RL = Ringer Locke solution (Armstrong *et al* 1953)

Tonicity Marked deviations from physiologically normal osmotic pressure also produce pain With decreasing concentrations of NaCl pain occurs at about 0 3 per cent NaCl and with increasing concentrations at about 3 per cent NaCl (Figure 3)

Acidity When acid solutions are applied to the blister area pain is not produced until the pH falls to about 3–3 5 (Figure 4)

Histamine The part played by histamine in tissue reactions and claims by Rosenthal that it is the mediator of cutaneous pain led us to test this substance on the blister area

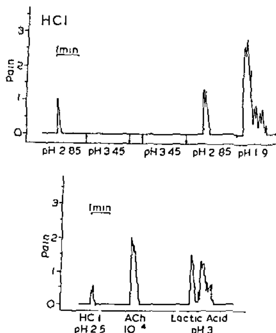


FIG 4 Production of pain in blister base by HCl and lactic acid Ordinate Pain Scale The acids were made up in Locke s solution to give isotonic solutions of varying pH (Armstrong *et al* 1953)

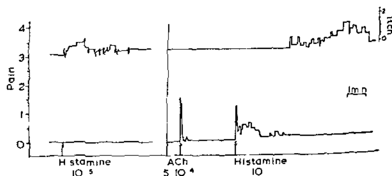


FIG 5 Production of pain and itch by histamine Upper record = itch lower record = pain Concentrations of histamine and ACh given in g /ml (Armstrong *et al* 1953)

Briefly, we have found that in concentrations of 10^{-5}g /ml or higher it can produce transient pain (which is sometimes severe with 10^{-3}g /ml concentration) the pain being followed by itching (Figure 5) Only itching occurred with concentrations from 10^{-5} to 10^{-7} g /ml Whenever there were sensory effects weal and flare were also seen

It is generally accepted that itching does not occur after the epidermis has been removed, so in our experiments it is likely that the itching induced by histamine is due to penetration of the substance into the normal skin surrounding the blister base Our results are compatible with Broadbent s (1953–5) suggestion that when it acts on intra epidermal endings histamine produces itching, but that when it penetrates to deeper nerve endings histamine can if the concentration is high enough produce pain We have been quite unable to confirm Rosenthal s claim that histamine produces cutaneous pain in a concentration of 10^{-18} g /ml

We found that a histamine releasing agent, compound 48/80 produced both pain and itching when applied to the blister base

It is difficult to know whether histamine release in skin or other tissues ever produces concentrations high enough to cause pain Certainly, the classical instance of *in vivo* release of histamine namely the urticarial weal is accompanied not by pain but by itching However one cannot exclude the possibility that with more severe tissue damage, high pain producing concentrations of histamine might be released at first but in these circumstances other pain producing agents might also be present as I shall describe in a moment

In the course of our earlier work we collected the fluid from the blister into a glass tuberculin syringe and in some experiments we reapplied this fluid to the exposed blister base, at first regarding this as a control application However we were surprised to find that on many occasions the blister fluid produced marked and prolonged pain after a latent period of 15–45 seconds On other occasions blister fluid did not produce pain and eventually we found that when the application was made within one hour of aspiration the fluid was almost

invariably painful whereas after one hour it was not (Figure 6) Thus the pain was produced by an unstable constituent of blister fluid

We naturally wondered what this substance could be and at first we thought it might be 5 hydroxytryptamine (serotonin,

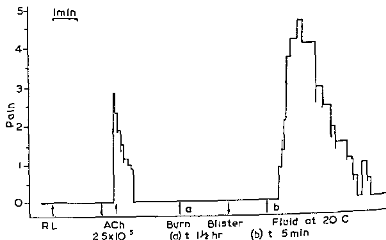


Fig 6 Production of pain by blister fluid Deep frozen blister fluid collected from a patient with a heat burn was later thawed and tested on the exposed base of a cantharidin blister in another person Ordinate = Pain Scale 5 = intolerable pain RL = Ringer Locke solution which produced no pain The area was sensitive to $2\ 5 \times 10^{-5}$g/ml ACh Burn blister fluid which had been melted for 1½ hours gave no pain (*a*) This was applied before the highly active sample which was melted only 5 minutes before application (*b*) so that the absence of response at 1½ hours could not be attributed to induced refractoriness of the blister area (Armstrong *et al* 1957)

5 HT) We therefore tested blood serum which was known to contain 5 hydroxytryptamine and found that this produced pain similar to that evoked by blister fluid, plasma being negative in these experiments (Figure 7*a*) Platelet extracts also produced pain By this time we had obtained synthetic 5 HT and we found that this produced the same type of pain as blister fluid and serum 5 HT being effective in concentrations of $10^{-6} - 10^{-7}$ g /ml and occasionally at 10^{-8} g /ml (Figure 7*b*)

Now since 5 HT is known to be very active in causing contraction of the isolated rat uterus we applied blister fluid to this

preparation and as expected observed uterine contractions (Figure 8) However, when we went into the matter further we found several respects in which the active substance in blister fluid differs from 5 HT For example the substance in blister

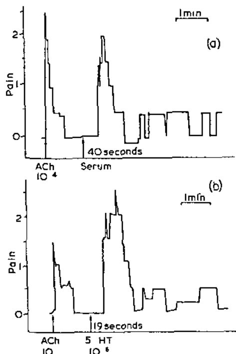


FIG 7 Pain responses in blister area to serum and 5 hydroxy tryptamine (5 HT Serotonin) (*a*) response to serum (*b*) response to 5 HT in a concentration of 10^{-6}g/ml (Armstrong *et al* 1953)

fluid is unstable but 5 HT is stable in this fluid, the action of 5 HT on the rat uterus is antagonized by dihydroergotamine and lysergic acid diethylamide that of blister fluid is not and acetone which extracts 5 HT from protein rich fluids precipitates the activity of blister fluid with the proteins

Thus we had found that both serum and blister fluid contained pain producing agents, that in serum being a stable compound of known structure namely 5 HT that in blister fluid

being an unstable substance of unknown composition We then found that the activity of blister fluid could be preserved by deep freezing which meant that we could keep such fluids until it was convenient to test them

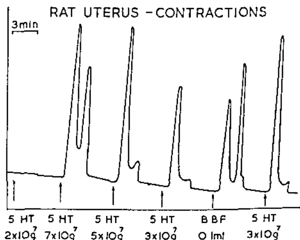


Fig 8 Contractions of isolated rat uterus in response to 5 HT and burn blister fluid (B B F) Volume of bath = 30 ml (Armstrong *et al* 1957)

We still had to consider the question—why is a blister not painful? For that is the case To study this point we tested fluid immediately after withdrawal from a blister and found that it did not cause either pain or uterine contraction but that after a short while it did both (Figure 9) Thus the pain producing, uterus stimulating activity developed only after collection of the blister fluid into the glass syringe Blister fluid was inactive *in situ*, and was activated after a few minutes in a glass syringe, the activity reaching a peak after 5–10 minutes and then decaying

We soon found that other inflammatory exudates such as a pleural effusion and joint fluid from patients with rheumatoid arthritis behaved just like blister fluid (Figure 10 Argent, Armstrong Jepson Keele and Phillips, 1954) Finally we found (Armstrong Jepson Keele and Stewart, 1957) that blood plasma responded in the same way (Figure 11) (Our original negative responses to plasma were due to testing more than one

hour after collection of the blood when the pain producing substance had come and gone)

Thus it seems likely that the pain producing activity of inflammatory exudates is due to their derivation from plasma in which the fundamental process of activation by glass can be much more conveniently studied Heparin does not affect

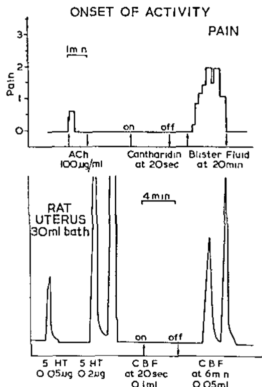


FIG 9 Development of pain producing and uterus stimulating activity in blister fluid on withdrawal into a glass syringe Upper record = pain lower record = contractions of rat uterus

The subject had two cantharidin blisters one on each forearm One blister was opened and prepared for testing Fluid from the other blister was removed and tested after 20 seconds on the blister base and rat uterus No activity was recorded After 6 minutes half the previous volume of cantharidin blister fluid (CBF) produced marked contraction of the uterus and after 20 minutes the aspirated blister fluid produced pain (Armstrong *et al* 1957)

this activation of plasma When blood is collected into a siliconed syringe, transferred to polythene centrifuge tubes containing heparin, and spun at 4000 rev /minute for 15–20 minutes, the upper half of the plasma contains no red cells or leucocytes

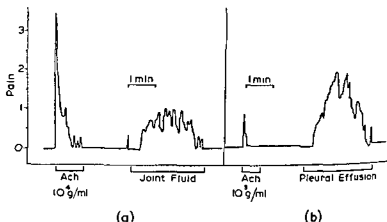


Fig 10 Pain responses in blister area to applications of (*a*) joint fluid from a patient with rheumatoid arthritis and (*b*) a pleural effusion of inflammatory type (Armstrong *et al* 1957)

and very few platelets Such plasma does not cause pain or contract the rat uterus (we call it preactive plasma), but when transferred to glass it is activated exactly as described for the exudates Thus virtually cell free plasma can form a powerful pain producing substance when brought into contact with glass

And now for the nature of this substance Using ultrafiltrates and aqueous and alcoholic extracts of active plasma, we have found that from its physical, chemical and biological properties the pain producing substance (PPS for short) appears to be a polypeptide closely resembling the substance bradykinin described by Rocha e Silva Beraldo and Rosenfeld (1949) (Figure 12) Bradykinin is made by the action of trypsin or proteolytic snake venoms on plasma α_2 globulins Plasma PPS resembles bradykinin in that it or something very similar, can be formed by the addition of trypsin to preactive plasma, the proteolytic enzyme of plasma called plasmin or fibrinolysin, acts in the same way Activation of plasma by contact with glass or by

addition of trypsin or plasmin is inhibited by the soya bean trypsin inhibitor

All these findings suggest that the fundamental process in the formation of PPS is activation of a proteolytic enzyme in plasma

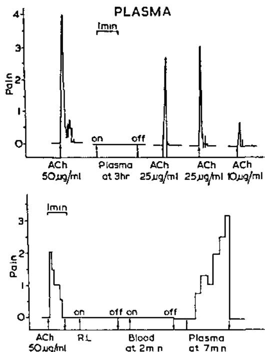


FIG 11 Production of pain by plasma on blister area The upper record shows absence of pain when plasma from heparinized blood was applied 3 hours after aspiration into a glass syringe The lower record shows absence of pain when heparinized blood was applied 2 minutes after aspiration the blood was centrifuged and severe pain was produced by application of the plasma at 7 minutes after aspiration

The rapid decay of the substance is presumably due to the presence of a peptidase in plasma (Armstrong *et al* 1957)

Now, before considering the possible roles of these various chemical agents in the production of the pain of tissue injury, I should like to refer to the receptors upon which they act The sensitivity of the cutaneous pain receptors to ACh (10^{-8}g /ml)

is comparable to that of the motor end plate, and it is interesting that the pain producing action of ACh is rather specifically antagonized by tubocurarine and decamethonium A further point do different substances act on different receptors? We

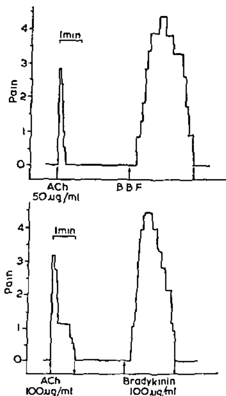


Fig 12 Production of pain by burn blister fluid and by bradykinin on the blister area

have no direct evidence on this point, but most subjects have little doubt that the immediate smarting pain of ACh is different in character from the delayed burning, pricking pain of 5 HT and plasma PPS The former feels more 'superficial than the latter and it may well be that ACh acts directly on the nerve endings exposed on the blister base and that 5 HT and PPS act at a deeper level—hence their long latent period

Chemical Agents in Pathological Pain

Next I shall consider how far these various chemicals can account for pathological pain It seems to me unlikely that there is a single chemical mediator for all types of pain, and comparisons with the role of ACh or noradrenaline as trans mitters, or of histamine in anaphylactic shock are over simplifications Good examples of extrinsic pain producing mixtures are provided by nettle stings and insect venoms

Nettle Stings Emmelin and Feldberg (1947) showed that nettle stings contain high concentrations of ACh and histamine which could certainly produce the initial stinging pain The itching and triple response would of course be due to the hista mine

Recently Collier and Chesher (1956) have shown that nettle stings also contain 5 HT in amounts which could produce pain of longer duration It is also intriguing to hear that Brittain and Collier (1957) have found an antagonist to 5 HT in dock leaves

But even with three pain producing agents it is still difficult to explain why the pain following nettle stings may last for some hours It seems to me possible that a proteolytic enzyme called solanain, which occurs in nettle stings might contribute to the picture by forming pain producing polypeptides from plasma and tissue proteins

Wasp Venom Jaques and Schachter (1954) have shown that wasp venom contains in addition to high concentrations of histamine and 5 HT a pain producing polypeptide named kinin There is no difficulty here in accounting for both pain and local tissue reactions in terms of these substances

Among other venoms, *bee* venom contains much histamine but no 5 HT and *scorpion* venom contains enough 5 HT to produce pain (Adam and Weiss 1956)

Now I would like to discuss the possible roles of the *intrinsic* pain producing factors This discussion is necessarily specula tive as there are no technical means of measuring local release of these substances while pain is occurring To me it seems most likely that pain resulting from tissue damage could be due to

histamine (in high concentration), 5 HT and the plasma PPS, though in the special case of peptic ulcer HCl can play a part, and it would be difficult to exclude the contribution of released potassium from damaged cells

Histamine would be liberated from damaged mast cells, and perhaps 5 HT as well, but release of the latter could arise in another way Trauma damages capillary endothelium which then becomes sticky Leucocytes and platelets adhere to the damaged endothelium, the platelets disintegrate and release 5 HT in high local concentration If blood escapes into the tissues, release of 5 HT from platelets would also occur The work of Beloff and Peters (1945) shows that tissue damage releases an active proteolytic enzyme in skin and if such an enzyme acted on plasma it could form the pain producing polypeptide Tissue cells also contain an activator of the proteolytic system of plasma

It is interesting to note that histamine, 5 HT and the plasma PPS form a mixture almost identical with that found in wasp venom It is as if such an unholy trinity had been evolved by Nature as the most reliable recipe for pain!

And now a few words about the possible relation of plasma PPS to blood clotting Macfarlane (1956) has long regarded blood clotting as just one of many reactions of blood to injury so it is of interest to consider whether the formation of plasma PPS is in any way related to clotting When blood is put into glass PPS formation occurs before visible clotting Anticoagulants such as heparin, citrate and oxalate do not prevent the formation of PPS by glass so if the two processes are related it would appear to be at some early phase of thromboplastin formation which does not depend on the presence of ionic calcium Other reactions of blood to injury could include activation of the permeability factor $G_2\alpha$, isolated from guinea pig plasma by Miles and his colleagues at the Lister Institute, as well as liberation of histamine, 5 HT and adenosine triphosphate from platelets (Humphrey and Jaques, 1955, Born, 1956)

Thus in trying to explain responses to injury we are faced with the possibilities that tissue cells, blood cells or plasma can

all release or form active agents which could contribute to the pain In such a case as the delayed pain after a burn (see Figure 1) histamine, 5 HT and plasma PPS could all be implicated The 5 HT might well come from platelets as a direct result of the heat, since we have found that heating platelets to 70–75 C for a few seconds releases this substance Dr Armstrong has found in skin extracts an additional pain producing agent which has not yet been identified but it is not histamine, 5 HT or a polypeptide In longer lasting 'spontaneous pain histamine and 5 HT would be less likely excitants because release of histamine tends to be explosive and because pain nerve endings become refractory to 5 HT Formation of PPS, on the other hand, could continue so long as there is proteolytic activity in the damaged tissues It seems likely that the continued presence of relatively low concentrations of histamine 5 HT and PPS could produce hyperalgesia and an additional factor which could increase the sensitivity of pain nerve endings is ischaemia of the nerves resulting from increased tension in an inflamed area

There is a little more to say about some of the conditions discussed earlier on On the subject of the pain of *peptic ulcer*, I agree not only with Palmer and with Bonney and Pickering that acid gastric juice can produce pain but also with Kinsella (1953) and with Smith (1955) that it is the state of inflammation in the ulcer which determines the pain response If hyperalgesia is marked any stimulus will provoke pain

With respect to *ischaemic muscular pain* I have no fresh suggestions to make about the nature of Lewis s factor P but I would like to make one point about the pain of coronary occlusion After the occlusion the affected heart muscle of course contracts ischaemically but it must be remembered that the infarct is an acute inflammatory reaction Thus in addition to factor P , I would expect 5 HT and plasma PPS to contribute to the pain of coronary thrombosis

A few words more about *migraine* In addition to distension of the temporal artery Ostfeld and Wolff (1955) have shown that there is oedema and a reduced threshold to deep pain on the affected side which they suggest is due to plasma PPS Thus the

headache results from stimulation of hypersensitive pain nerve endings '

"It would be quite wrong in a talk about pain to omit some reference to *toothache* and I would like to mention briefly some work now being done by Anderson and Howard at Guy's Hospital They have shown that the exposed dentine is generally insensitive to chemical excitants of pain although sensitive enough to temperature changes particularly to cold, and to pressure For toothache the most important factor would appear to be the inflammatory changes in the pulp, with the ensuing hyperalgesia which can rise to pain at the slightest external stimulus or even from the pressure exerted by structures within the diseased tooth on the nerve endings in the pulp '

Next a few words about possible antagonists to the pain producing substances which I have mentioned *Local anaesthetics* will of course prevent their action by blocking the pain nerve fibres peripherally, but we have not come across any specific antagonists to plasma PPS Salicylates have a slight action but centrally acting drugs like morphine show no peripheral antagonism to PPS Cortisone does not prevent activation of PPS by glass, but we have not studied its antagonism to the action on the blister area It is difficult to study antagonists to 5 HT on the blister area owing to the refractory state which repeated applications of 5 HT induce

It is quite clear that there are many questions still to be answered, particularly in regard to liberation and activation of pain producing substances *in vivo* I hope that our experimental results and the speculations which have developed from them will stimulate others to an interest in a subject which we have found very fascinating

REFERENCES

ADAM K R and WEISS C (1956) *Nature* **178** 421

ARGENT D E ARMSTRONG D JEPSON J B KEELE C A and PHILLIPS L A (1954) *J Physiol* **124** 18P

ARMSTRONG D DRY R L M KEELE C A and MARKHAM J W (1953) *J Physiol* **120** 326

Armstrong D Jepson J B Keele C A and Stewart J W (1957) *J Physiol* **135** 350

Beloff A and Peters R A (1945) *J Physiol* **103** 461

Bonney G L W and Pickering G W (1946–8) *Clin Sci* **6** 63

Born G V R (1956) *J Physiol* **133** 61P

Brittain R T and Collier H O J (1957) *J Physiol* **135** 58P

Broadbent J L (1953) *Brit J Pharmacol* **8** 263

Broadbent J L (1955) *Brit J Pharmacol* **10** 183

Collier H O J and Chesher G B (1956) *Brit J Pharmacol* **11** 186

Doupe J Cullen C H and Chance G A (1944) *J Neurol Neurosurg Psychiat* **7** 33

Emmelin N and Feldberg W (1947) *J Physiol* **106** 440

Hardy J D Wolff H G and Goodell H (1952) *Pain Sensations and Reactions* Baillière Tindall and Cox London

Humphreys J H and Jaques R (1955) *J Physiol* **128** 9

Jaques R and Schachter M (1954) *Brit J Pharmacol* **9** 53

Kinsella V J (1953) *Lancet* **ii** 353

Lewis T (1942) *Pain* The Macmillan Company New York

Lewis T Pickering G W and Rothschild P (1931) *Heart* **15** 359

Macfarlane R G (1956) *Physiol Rev* **36** 479

MacWilliam J A and Webster W J (1923) *Brit med J* **i** 51

Nathan P W (1953) *J Neurol Neurosurg Psychiat* **16** 144

Ostfeld A M and Wolff H G (1955) *Arch Neurol Psychiat* **74** 131

Palmer W L (1926) *Arch intern Med* **38** 694

Pickering G W (1939) *Brit med J* **i** 907

Pickering G W and Hess W (1933) *Clin Sci* **1** 77

Richards R L (1951) *Clin Sci* **14** 76

Rocha e Silva M Beraldo W T and Rosenfeld G (1949) *Amer J Physiol* **156** 261

Rosenthal S R and Sonnenschein R R (1946) *Amer J Physiol* **155** 196

Scott R B and Warin R P (1946–8) *Clin Sci* **6** 51

Smith A W M (1955) *Quart J Med* **24** 393

Walshe F M R (1948) *Critical Studies in Neurology* p 34 E & S Livingstone Ltd Edinburgh

Weddell G Palmer E and Pallie W (1955) *Biol Rev* **30** 159

Wolff H G (1948) *Headache and other Head Pain* Oxford University Press New York

X

Neurochemistry

HENRY McILWAIN

WHEN I was invited to lecture on a neurochemical topic I provisionally accepted the simple title of Neurochemistry, thinking that if necessary some one aspect of this large subject could be chosen at a later date However, in the meanwhile my curiosity was aroused by the increasing use of the term neurochemistry and it appeared valuable to discover exactly what activities this newly-proliferated name was covering So I propose to explore this in a way which exhibits the development of the subject and the contribution which it makes to the scientific basis of medicine which gives some indications of our own current work and displays a little the relationship between the biochemical specialist and his parent subject

NEUROCHEMICAL LITERATURE

At least in the past decade, the term neurochemistry has been used to include all aspects of chemistry biochemistry and chemical treatment related to the nervous system and to the neurological and psychiatric half of medicine (Elliott 1947)[1] These medical specialities have been especially prolific in the number of compound terms which they have engendered Neuroanatomy neurophysiology psychopathology have become current while hepatophysiology or myopathology have not, perhaps, therefore, a neurochemistry was to be expected before a hepatochemistry or a myochemistry If there is a basis for this in the subjects themselves it presumably lies in the

[1] The term *Nervenchemie* was employed in the last century to indicate more literally the chemistry of nervous tissues

outstanding contribution which the nervous system makes to the dominance of mammals and primarily of man This significance of the neurological sciences has long been felt, in 1833 J P Couerbe presented to the Académie des Sciences (Paris) a paper which dealt primarily with the fractionation of cerebral lipids but which he prefaced by making a plea for the fixing of the attention of all our scientists on the nervous system, and in which he posed questions on the nature of man and his intelligence These questions he tackled by the neurochemical methods of his time he attacked several pounds of ox brain with ethanol and ether with the object of separating and chemically characterizing its constituents He applied this method not only to the brain of animals but also to human brain, and thought that he found chemical differences between specimens from the normal the insane, and the idiot

This type of study constitutes the beginning of neurochemistry contemporary with the beginning of organic chemistry and biochemistry themselves The theme of separating and identifying cerebral constituents especially lipids was multiplied, but not diversified by J L W Thudichum whose *Chemical Constitution of the Brain* (1884 1901) can be taken as typical of the picture of neurochemistry at the turn of the century From then almost until the present decade the physiological chemistry of the nervous system tended to attract much less attention than its electrophysiology but Winterstein s (1929) monographs are an exception When Page (1937) wrote his book, the biochemistry of the nervous system itself was established and several current themes are included So we come to the 1950 s with books by Himwich (1951) and Mc Ilwain (1955), and four collections of papers specifically on the subject of neurochemistry two symposia (1955 1957), and the others edited by Elliott Page and Quastel (1955) and Korey and Nurnberger (1956) It is these four collections and the chemical papers appearing in current journals of neurology and psychiatry which I will take as typical of the present state of neurochemistry

Many neurochemical studies of course appear in biological biochemical physiological and pharmacological journals

and have begun to appear in the new *Journal of Neurochemistry* These are at present probably exceeded by the numbers appearing in medical journals, some twenty neurological and psychiatric journals carry appreciable numbers of such papers For the present survey I have chosen those appearing in a few unselected numbers of one British journal, the *Journal of Neurology, Neurosurgery and Psychiatry* and one American, the *Archives of Neurology and Psychiatry* It is remarkable that chemical matters constitute a dominant part of 35 to 40 per cent of all the papers in these numbers

Starting with this sample of 82 papers from journals, 50 from collected papers, and 98 from symposia, the classification of Tables 1 and 2 emerges This first divides the subject into two approximately equal parts in the first (Table 1), the nervous system or preparations from it are studied largely independently of the rest of the body In the second group (Table 2) the dominating theme is the relationship of the nervous system with the body as a whole The two aspects of course overlap and are interdependent

In all these different sources of neurochemical data, studies concerning the central nervous system preponderate This is due, apparently to a combination of causes to which contribute its importance and its availablity, especially the large lump the brain, readily recognizable by biochemists although they may have to be asked how much of the cerebellum or brain stem they have included in their samples

Direct Studies of the Nervous System Itself

In this and the following section, paragraphs are numbered according to the headings of Tables 1 and 2 Of the investigations quoted many have been carried out at the Maudsley laboratories and to these P Greengard, P J Heald H Martin R Rodnight and G H Sloane Stanley have contributed

1 As already indicated the first fundamental chemical study of the nervous system was of its composition Table 1 bears witness to the way in which this type of investigation has continued New major constituents are still being found among the lipids, and these and the proteins of the nervous system have

yet to be examined from many points of view One, now increasingly studied, is their localization in different parts of the nervous system and also in different subcellular components ⁄

TABLE 1 Subjects included in Neurochemical Publications

Subject	Thudichum (1901)	Winterstein (1929)	Page (1937)	Collected papers (1955 1956)	Symposia (1955 1957)	*J Neurol Neurosurg Psychiat* (1954 1955) *Arch Neurol Psychiat* (72 1954 73 1955)
1 *Chemical composition*						
Relatively stable	*	*	*	*****	**********	******
Labile	0	0	0	*	*	*
In relation to cell structure	0	0	0	0	*******	*
2 *Metabolism*						
Lipid	0	*	(*)	**	******	*
Carbohydrate respiration	0	*	*	(*)(*)**	* ***	***
Proteins and amino acids	0	*	(*)	***	** ***	**
Nucleic acids	0	0	0	**	***	
Phoshorylation	0	0	0	(*)	*******	*
Of coenzymes *et al*	0	0	0	0	**	*
3 *Metabolism and transmission or electrical phenomena*						
Acetylcholine and enzymes	0	0	*		* *	**
Noradrenaline and adrenaline	0	0	0	*	*	**
Other probable transmitters	0	0	0	0	**	0
Electrolyte movements	0	0	*	*	*	*
Electrical phenomena and metabolism	0		*	0	***	******

The * signs refer to the number of papers on the subject indicated or to the inclusion of major accounts of the particular subject in books by the authors named (*) signs imply the subject is discussed largely on the basis of data from parts of the body other than the nervous system

But new, purely chemical discoveries remain to be made even in kephalin, the lipid fraction which some ninety years ago took its name from the brain discovery of an inositol containing

compound as a major constituent is recent Still more recent is the finding, in the Maudsley laboratories, that the inositol lipid undergoes enzymic change with cerebral preparations at rates 100 times greater than those previously recognized in cerebral lipids

2 This illustrates how in earlier days a substance attracted attention because it was present naturally in large amounts, now it tends to attract attention when it undergoes rapid metabolic change This gives a different outlook and direction to the study of the chemical composition of an organ by analysing it before and after the induction of brief, rapid changes in functional activity indication is obtained of the metabolic changes which are closely associated with such activity Excit able tissues are pre eminently suited for studies of this type, which reached a high level in relation to peripheral nerve when von Muralt (1946) working at intervals of milliseconds only, tried to 'freeze' a nerve impulse as it passed along an isolated nerve A much wider series of changes has been studied in the brain, fixing the tissue for analysis at intervals of one or a few seconds after electrical excitation Within this brief period a number of cerebral constituents have been found to increase or decrease in level These include substances such as acetyl choline and also many which are concerned in energy yielding and transferring processes glucose glycogen, lactate, and creatine, adenosine, and inorganic phosphates It is of course necessary in an organ which is to guide bodily action that response should be rapid Response is usually thought of in electrical terms and the study of chemical composition shows how extremely closely metabolic response in the brain is associated with electrical phenomena This raises one of the most fundamental metabolic problems regarding the nervous system namely, how its metabolism is controlled For the increased metabolism associated with activity can briefly, rise to many times the resting level glycolysis, for example ten fold

The actual processes of carbohydrate and related metabolism, when studied in detail, have important quantitative aspects which are specific to the nervous system, including these prob lems of speed and integration with tissue function This has been

documented elsewhere (McIlwain, 1955) In a qualitative sense the processes are often similar to those of muscle and other tissues and indeed the primary description of certain aspects of carbohydrate metabolism came from studies with cerebral tissues This was especially so in relation to the role of co carboxylase and thiol derivatives in pyruvate metabolism In contrast however the energy consuming processes appear to be more specific to nervous tissues Any discussion of these processes and of control of metabolism in nervous tissues must lead back to the central problem of how nervous transmission occurs Here the available information tantalizingly awaits chemical linkage Material, ionic exchange is the basis of the nerve impulse In the spike potential, exchange is too rapid to be enzymic but probably requires carrier molecules and if so these should be present in appreciable quantity and isolable So also should be more components of the systems expending metabolically derived energy in maintaining the ionic differences on which the spike potential depends Here the contribution of the Maudsley Laboratories has been to try to follow the fate of the additional phosphate which leaves energy rich compounds in a brief period of seconds when cerebral tissues are provoked to increased activity by electrical pulses For such studies it is important that means have been found of exciting electrically isolated mammalian cerebral tissues for the high metabolic turnover of such tissues and the greater quantity available give many technical advantages over studies with peripheral nerve When radioactive phosphate is used, a few seconds electrical excitation is in fact found to increase greatly the labelled phosphate transferred from the labile phosphates to other nucleotides and to phosphoprotein These groups of compounds therefore deserve further study Between the energy rich compounds and the ion movement which they support must come defined chemical interaction with other substances as part of a membrane a protein as in muscle or hypothetical molecules which thereby can bind release exchange or displace sodium or potassium ions

3 The phenomena which have been described are among the most fundamental in neurochemistry—those concerned with

ion movement They also form its major unsolved problems the mechanism of ion movement during the nerve impulse, and the mechanisms by which the tissue maintains its differential ion concentrations inside and outside the nerve fibre These constitute the centre of neurochemistry much as the properties and interactions of atoms and molecules make the basis of chemistry, and those of enzymes and metabolites the centre of biochemistry Granted the chemical basis of the firing of the individual neurone, and that transmission of its impulse along the nerve fibre is such a central feature of the subject, why does it not attract more work? (I am not speaking here of transmission between cells) Presumably because it is an inherently difficult subject, and because other aspects of neurochemistry are also important and often seem nearer to clinical application For one could not expect neurochemistry to become a deductive science if chemical mechanisms of transmission were known, the situation is more comparable to knowledge of molecular structure in chemistry which has rather, made conjecture and anticipation more accurate and varied On the other hand one cannot expect difficult, fundamental problems to be solved without considerable deliberation, time, effort and resources Probably only when they are in the stage of formulation can such problems best be approached sidelong, when clearly perceived one judges the best attack to be a direct one Success in understanding the chemistry of impulse transmission could be very important, especially in bringing increased understanding of the make up of the nervous system as a whole, and of the way in which it is affected by normal bodily events and by drugs

The phenomena which have just been described concern the individual nerve cell and are common to almost all parts of the nervous system of all species studied (though probably here, as elsewhere, comparative biochemistry has much to contribute) Chemical diversity and specificity are however extremely prominent in the nervous system as is made manifest by the very names adopted in classifying neurones as adrenergic or cholinergic We here come to a second category of chemical factors related to transmission in the nervous system largely concerning transmission between cells These substances are

involved in quantities small in comparison with the changes in sodium and potassium ions already described but they are of great qualitative importance They have proved more tractable to chemical study than the systems concerned in ion transport so that much is now known of the formation and breakdown of acetylcholine and the way in which the functioning of the nervous system is altered when such metabolism is disturbed Nevertheless in most parts of the nervous system the synthesis and breakdown of known transmitter substances by known routes can involve only a small proportion of their energy turn over Most presumably goes, by routes already adumbrated to support ion differences and the readiness to transmit The transmitters one regards rather as triggers, making feasible, for example the suggestions of abnormality in the metabolism or binding of acetylcholine in areas of abnormal epileptic dis charge But in the central nervous system it is especially promising to see that exploratory work of the last few years is in dicating still further chemical diversity among transmitter sub stances, suggesting for example serotonin and aminobutyric acids *to be involved in transmission phenomena*

The Nervous System and the Rest of the Body

4 With problems of intercellular transmission one is clearly approaching the linkage not only between different parts of the nervous system but also between the nervous system and the rest of the body This latter relationship involves in one respect a great elaboration of fine anatomical detail, for example at nerve endings at glands or muscles in other respects it involves much chemical detail as in the maintenance of the nervous system itself and in the maintenance of homeostasis when the nervous system responds to chemical stimuli and elaborates agents making humoral connection with the rest of the body

Relationships between the brain and the body have long been the subject of speculation nearly a century ago Henry Mauds ley was among the most insistent in emphasizing that the brain *required a large supply of blood and that it was highly sensitive* to materials supplied by the blood (McIlwain 1955) Though he was correctly oriented in these respects Maudsley was not a

laboratory worker For all those who have speculated on the interactions between the brain and the body, only a handful have investigated the interchange at its major route by analysing the blood as it enters and as it leaves the brain Of the major organs of the body, the brain is the one most immediately dependent on second to second material supplies from the blood This is obvious to all from the promptness with which unconsciousness comes when the blood supply to the head is stopped I find it valuable to emphasize this at the beginning of biochemical lectures in the University course in psychological medicine, indicating that in studying the mind, one is studying the bodily activity most critically and continuously dependent on material, chemical factors We shall see further examples of this shortly, but for the present I wish to describe the evidence obtained by analysis of arterial and cerebral venous blood

Such studies began in the 1930s to give a consistent picture showing the importance of carbohydrate metabolism in maintaining cerebral activities The balance sheet of glucose and oxygen utilized, and carbon dioxide, lactate and pyruvate *formed was apparently self contained and consistent* with glucose being the only major substrate from the blood stream oxidized by the brain This was true although separated portions of the brain could actively oxidize a wide variety of added substances, for example glutamic acid, yet arterial and venous glutamic acid differ relatively little, and such difference as exists consists of the formation of the amide glutamine

Obviously other metabolic relationships exist between the brain and the rest of the body Vitamins and trace metals must be acquired, but these can be conceived as relatively slow processes Perfusion studies, however have suggested more diversified chemical exchange at the brain for simple glucose salines, or even quite complex ones with proteins or red blood cells added, do not maintain normal cerebral activities unless, curiously, the liver is included in the perfusion system, when some approach to normal is found

5 Description so far in this lecture has mainly concerned the normal adult nervous system Very extensive chemical investigation has understandably, been made of the nervous system

TABLE 2 Subjects included in Neurochemical Publications (continued from Table 1)

Subject chemical aspects of	Thudichum (1901)	Winterstein (1929)	Page (1937)	Collected papers 1955 1956)	Symposia (1955 1957)	*J Neurol Neurosurg Psychiat* (1954 1955) *Arch Neurol Psychiat* (72 1954 73 1955)
4 *Interaction between brain and body*						
Measured by blood analysis	0	0	*	**	**	* 12
Perfusion and related methods	0	*	*	0	*	0
Nutrition	0	*	*	*	0	**
Blood brain CSF barriers	0	*	0	*	*****	** ***
Hormonal	0	0	0	*	***	* 15
Measured by urinary excretion	0	0	0	0	**	******
5 *Physiological and pathological variables*						
Genetic	0	0	0	*	*	*
Developmental and degenerative changes	0		*	**	* 34	***
Supply of metabolites	0	*	*	*	* *	*********
Convulsive conditions	0	0	0	**	**	********
Neurological disorder	0	0	*	**	**	* 26
Mental and emotional disorder and defect	0	0	0	*	**	* 25
6 *Pharmacological and toxic agents*						
Neurotoxins and venoms	0	0	0	*	0	***
Metals metabolic inhibitors and antagonists	0	*	*	*	**	* * ******
General depressants	0	*	*	*	*	***
Specific depressants	0		0	*	0	** ** *
Drugs mainly excitant	0	*	0	*	0	****
Others including side effect of drugs	0	0	0	0	0	** **

Nomenclature as Table 1 except that when many papers appeared in a given subject their number is given by figures

N

during development and degeneration and also of the brain and the rest of the body in nervous, mental and emotional disorders The first neurochemical symposium concerned the development of the nervous system, and gave an impressive picture of the march of chemical, electrical and anatomical events while the brain is taking over its adult functions We for example showed how, in completely separated cerebral tissues from animals of different ages, their metabolic characteristics changed so that just ahead of the growth of electrical activity in the brain came the ability to obtain metabolic response to applied electrical pulses, the brain was being prepared metabolically for its new functioning just before such functioning began Perhaps the most important general fact in this connection is that the greater part of the nervous system is formed actually during its functioning Even in the brain, electrical activity begins when the organ is only about one fifth of its adult weight Growth during functioning carrying the possibility of moulding to functional need, is a commonplace in most organs of the body—in the muscles for example—but can be realized also as of very great importance in the brain It makes very real the possibility during the first few years of human life, of building a representation of the environment into the individual

Exploration of chemical factors in the nervous system s reaction to physiological extremes and to pathological change has a long history For a brief exposition I cannot do better than remind you of Barcroft s half impertinent questioning of Claude Bernard s famous aphorism Bernard wrote La fixite du milieu internal, est la condition de la vie libre and Barcroft (1934) bluntly queried What is this free life of which the fixity of the internal medium is the condition? His answer is exhibited in Table 3 All these changes in the level of bodily constituents affected first the nervous system and especially the brain The free life disturbed or disrupted by these extremes, Barcroft concluded to be the life of the intellect of normal mental and emotional behaviour, since this was the first to suffer when bodily chemistry was disturbed in all the various ways Examples could now be increased greatly those referred to in the collected

papers include thiamine pyridoxine, nicotinic acid vitamin B_{12} copper salts and galactose

6 I have left until last a description of how the study of pharmacological agents fits into neurochemistry for I feel it is

TABLE 3 Effects of changed concentrations of bodily constituents (Largely from J Barcroft 1934 see Himwich 1951 and McIlwain 1955 1957)

Constituent	First result of change in level	
	Decrease	Increase
Water	Weakness minimizes epileptic tendency	Headache nausea dizziness asthenia incoordination
Hydrogen ion (pH)	Headache	Coma
Oxygen	Headache lassitude poor performance in mental tests	Convulsions
Glucose	Changes in sensation and behaviour coma	Slight narcotic action
Sodium salts	Fever	Reflex irritability weakness paresis
Calcium salts	Nervous twitchings convulsions	Apathy drowsines verging on coma general atonia

at present one of the least satisfactorily accommodated yet most important The reason for the importance is two-fold First over half the existing agents of the pharmacopoea act primarily on the nervous system Secondly when neurochemical investigations lead to new therapeutic applications they naturally lead most directly to chemical means of therapy Neurochemical studies may of course have other routes of application as diagnostic tests or in pointing to dietary measures But the major triumph of chemistry in its application to disease and discomfort involving the nervou system has been in the production of synthetic drugs from the well established local anaesthetics analgesics sedatives curare like substances depressants anti convulsants and many other classes to the excitants and depres sants of the central nervous system which are now being exa mined in relation to mental illness The ability to produce these drugs must be contrasted with our ignorance of how most of

them act Of those named, only the curare like substances can be said to be reasonably understood in action Ignorance regarding the others underlines our lack of knowledge of chemical factors in nervous transmission Only recently, study of the electrically stimulated metabolism of cerebral tissues and of ganglia has demonstrated metabolic effects of anticonvulsants or general depressants on those parts of the nervous system at which they have their primary actions, and in the low concentrations at which they do act

Unfortunately, actions of chemical substances on the nervous system are not always of medicinal value Rather, they may limit the employment of compounds synthesized, for example, as antimalarials or as local anaesthetics Many inorganic substances have their toxic effects also at the nervous system, as also do several toxic agents of plant and animal origin, and Table 2 bears witness to the continued preoccupation with such agents, as seen by the number of papers in current neurological and psychiatric literature There are not as yet a corresponding number of papers on such subjects in the neurochemical symposia or collected papers, which reflects the lack of understanding described earlier Clearly, in spite of protective barriers' the nervous system remains a major target for chemical substances of almost all sorts, introduced to the body I would like to point out a most interesting confirmation of this in the way in which the synthetic drugs discovered during the last century largely acted at the nervous system and to a great extent did not have the effects intended by the investigators who first studied them The actual sequence is shown in a simplified and schematic form in Table 4

It will be recalled that throughout the last century the most pressing medical problems were those of infectious disease Even Thudichum, in relating his study of the chemistry of the brain to practical problems, gave more prominence to cholera and typhoid fevers than to general paralysis of the insane, and more to this than to other forms of mental disease This preoccupation with infectious disease makes it understandable that the substances which found application as general anaesthetics were first studied in relation to consumption and other

diseases of the lungs The first antipyretics and analgesics came from agents which were intended to combat fevers and infection salicyclic acid and quinoline derivatives Salicyclic acid Kolbe found to prevent putrefaction *in vitro* and he therefore anticipated activity *in vivo* against infection This at first appeared

TABLE 4 Intention and discovery in the first synthetic drugs (From McIlwain 1957)

Group of substance and examples	First studied in relation to	Application found
Simple gaseous or volatile compounds N O ether	Diseases of the lungs	General depression
Simple aliphatic alcohols aldehydes paraldehyde chloral	Sedation	Sedation
Simple aromatic compounds aniline phenol salicylates	Infectious disease	Antipyresis mild analgesia local antisepsis
Heterocyclic compounds easily derived from simple aliphatic and aromatic compounds quinolines pyrazolones barbiturates	Infectious disease sedation	Antipyresis sedation

confirmed when he found it to lower body temperature, only later was the effect shown to be on the heat regulating centres of the medulla oblongata The quinoline and pyrazolone derivatives, made in imitation of quinine, proved to act in the same way and to have no more than symptomatic effect in fevers The table includes one category of substances which had the medicinal effect which was intended by those who investigated them This is the group of sedatives or hypnotics where a central action was desired and obtained The table may be said in brief to summarize an experiment which lasted almost a century the exposure of man and other animals to representative compounds of almost all the main groups of organic chemicals as successively organic synthesis made them available Thus Table 4 *exhibits the sensitivity of the nervous system to synthetic* substances much as Table 3 giving Barcroft s experiments exhibited the sensitivity of the brain to naturally-occurring compounds

The question of how synthetic drugs are produced is no more than touched on in Table 4, in its entirety it presents an enormous amount of chemical data regarding the nervous system Thus the table mentions briefly only a few classes of compounds Within these classes and many others, much effort has been expended in finding the structures optimal for different desired actions, and avoiding toxic effects How to produce the most valuable member of a series—in essence, the practice of chemotherapeutic research—was learned during the last century primarily in connection with agents affecting the nervous system This however is a long story which is told more fully elsewhere (McIlwain, 1957) I will note here only the fact that this development did take place largely in relation to the brain and that this reflects the significance of chemical factors in relation to the nervous system, in a way characteristically different from that of the rest of neurochemistry

Neurochemists

Finally, I would like to comment on the sort of people who are doing neurochemical work At present, they are of very many different backgrounds For while no one has been trained primarily as a neurochemist, chemical and biochemical workers are in continuous demand for studies in relation to the nervous system and its ailments or for research in laboratories which study the nervous system as part of physiology, endocrinology or pharmacology Taking as example those who have received training in our own laboratories, one finds that they have come from or gone to departments of biochemistry, medicinal chemistry chemical pathology pharmacology physiology, other medical specialities and the laboratories of general and mental hospitals

Present training is gained in an apprenticeship fashion, to a large extent by carrying out research Most established neurochemical workers have a double or triple qualification perhaps a primary degree in chemistry, biochemistry physiology or medicine, a higher degree in a complementary subject and subsequent training in work requiring knowledge of neuroanatomy, neurophysiology or pharmacology Probably training

in a subject such as neurochemistry will remain for many years a postgraduate speciality, but one feels that the time has come when it could be a primary postgraduate speciality certainly no narrower than crystallography, geophysics or microbiology Potentially, neurochemistry can be considered a more funda mental subject than is pharmacology as chemistry is more fundamental than medicinal chemistry

Where do present neurochemical workers receive their train ing? For work on chemical aspects only of the nervous system and in collaboration with other investigators some have come, without further special training from departments concerned with the chemistry of natural products as a whole Similarly with respect to purely metabolic work others have come from departments of biochemistry where the nervous system receives some comment as one of the twenty or thirty other special systems studied, which range from bacteria and plants to the organs of primitive and higher animals The wide background has obvious advantages but leaves the necessity for specifically neurochemical learning or training later This may come by collaborative work It is only during such work—and perhaps only incidentally during it—that major outstanding problems in neurochemistry as a whole are appreciated

A worker trained in general chemistry or biochemistry tends to see as central problems of chemistry or biochemistry of which the nervous system presents examples A chemical worker collaborating in a department of pathology or in a mental hospital is led to study aspects of the nervous system of current importance in pathology or in mental disease Each approach has great value but there is beyond them the neces sity to see neurochemistry as a whole so that the feasibility and value of the separate chemical viewpoints can be appraised There has for instance been trenchant criticism of abuse of chemical effort in studying mental and emotional disorders (Altschule 1953) Moreover much real progress in neuro-chemistry has come when the nervous system itself has been taken as central and different chemical aspects of its functioning have been co-ordinated when metabolism *in vivo* is compared with that *in vitro* when chemical aspects of the composition and

functioning of primitive nervous systems are compared with those more highly developed, when a series of potential therapeutic agents are studied metabolically and from the point of view of their distribution, both in the whole animal and in separated systems These studies usually require a wider range of techniques than are found in one laboratory Also, unless resources are adequate there is a tendency to study some convenient and already developed system rather than the most appropriate, the action of anticonvulsants on peripheral nerve, for example, or of agents of very various types, on the cerebral cortex of the rat

There are probably not more than four or five places in the world where studies have been made in each of the six major categories of Tables 1 and 2 At the Maudsley in the past eight years, both research and teaching have included all these divisions, but it has never been possible to maintain work in more than a few of the divisions at any one time Much has been carried out with grants on a yearly basis, afterwards the worker with his specialized knowledge has departed, often developing his part of the subject elsewhere, but unavailable for the collaboration between widely different branches of neurochemistry which has produced most results and is necessary for tackling major unsolved problems

Until the central problems of neurochemistry have been successfully tackled and we see more clearly how the nervous system utilizes its large energy supply in nervous transmission and in maintaining the system in a state of readiness to react and how the brain is moulded to an animal s experience, chemical aspects of most of the neurological sciences—and above all, material approaches to nervous mental and emotional diseases—remain as empirical as was organic chemistry before the advent of structural formulae

REFERENCES

ALTSCHULE M D (1953) *Bodily Physiology in Mental and Emotional Disorders* Grune and Stratton New York

BARCROFT J (1934) *Features in the Architecture of Physiological Function* Cambridge University Press

COLLECTED PAPERS (1955) *Neurochemistry* Ed Elliott Page and Quastel Thomas Illinois

COLLECTED PAPERS (1956) *Neurochemistry* Ed Korey and Nurnberger Hoeber New York

COURBE J P (1833) *Ann Chim Phys* **56** 160

ELLIOTT K A C (1947) *McGill Med J* **16** 293

HIMWICH H E (1951) *Brain Metabolism and Cerebral Disorders* Williams and Wilkins Baltimore

McILWAIN H (1955) *Maudsley Mott and Mann on the Chemical Physiology and Pathology of the Mind* Lewis London

McILWAIN H (1955) *Biochemistry and the Central Nervous System* Churchill London

McILWAIN H (1957) *Chemotherapy and the Central Nervous System* Churchill London

MURALT A v (1946) *Die Signalubermittlung im Nerven* Birkhauser Basel

PAGE I H (1937) *The Chemistry of the Brain* Thomas Illinois

SYMPOSIUM (1955) *First International Neurochemical Symposium* Ed Waelsch Academic Press New York

SYMPOSIUM (1957) *Second International Neurochemical Symposium* Ed Richer Pergamon Press London

THUDICHUM J L W (1901) *Die Chemische Konstitution des Gehirns* Pietzcker Tubingen

WINTERSTEIN H (1929) In *Bethe s Handbuch der Normalen und Pathologischen Physiologie* Springer Berlin

XI

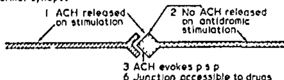
The Metabolism of Acetylcholine in Nervous Tissue

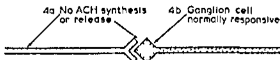
V P WHITTAKER

IT is now generally accepted that transmission at certain nerve endings involves the synthesis, release and destruction of acetylcholine The classical examples of this type of transmission are found within the sympathetic ganglion where a cholinergic pre ganglionic fibre synapses with a non cholinergic post ganglionic fibre and at the motor end plate where impulses are transmitted from the cholinergic nerve fibre to the end plate of the muscle The relatively isolated anatomical localization of these junctions has made their experimental investigation simpler than that of synapses in the central nervous system but there too there is good evidence though of a less direct kind for the existence of cholinergic transmission

The evidence that acetylcholine is the transmitting agent in sympathetic ganglia is summarized in Figure 1 and is as follows (1) stimulation of the pre ganglionic fibre releases acetylcholine in the ganglion whereas (2) antidromic stimulation of the ganglion cells does not (3) acetylcholine in small doses evokes the characteristic post synaptic potential (p s p) (4) the ability of the pre ganglionic fibre to conduct impulses to release acetylcholine and to synthesize acetylcholine *in vitro* disappear on degeneration (Banister and Scrase 1950 Hebb and Waites 1956), while the post synaptic membrane remains responsive to acetylcholine, (5) the ganglion cells cease to respond to acetylcholine if caused to degenerate while the pre ganglionic fibres retain their ability to synthesize and release acetylcholine (6) the post

synaptic region is accessible to, and its response to acetylcholine is powerfully modified by, certain drugs which do not affect peripheral conduction. Similar arguments apply, *mutatis mutandis*, to the motor end plate.





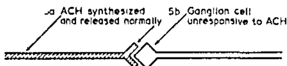


FIG. 1. Evidence for transmission by acetylcholine (ACH) at a sympathetic ganglion.

The success of acetylcholine as a transmitting agent depends upon its prompt removal from the site of action when its work of exciting the post-junctional membrane is complete. Though passive diffusion may under certain circumstances play a not inconsiderable part in this process (Ogston, 1950), it is powerfully reinforced under normal conditions by the presence in the junction of enzymes—the cholinesterases—which catalyse the hydrolysis of acetylcholine to the physiologically inactive acetic acid and choline. If the destruction of acetylcholine is interfered with—for example, by means of a cholinesterase inhibitor—prolonged depolarization of the junction will ensue, eventually passing over, in the case of the neuromuscular junction, into a non-depolarized type of blockade (Thesleff, 1955).

XI

The Metabolism of Acetylcholine in Nervous Tissue

V P WHITTAKER

It is now generally accepted that transmission at certain nerve endings involves the synthesis, release and destruction of acetylcholine The classical examples of this type of trans mission are found within the sympathetic ganglion where a cholinergic pre ganglionic fibre synapses with a non cholinergic post ganglionic fibre and at the motor end plate where impulses are transmitted from the cholinergic nerve fibre to the end plate of the muscle The relatively isolated anatomical localization of these junctions has made their experimental investigation simpler than that of synapses in the central nervous system but there too there is good evidence though of a less direct kind, for the existence of cholinergic transmission

The evidence that acetylcholine is the transmitting agent in sympathetic ganglia is summarized in Figure 1 and is as follows (1) stimulation of the pre ganglionic fibre releases acetylcholine in the ganglion whereas (2) antidromic stimulation of the gan glion cells does not (3) acetylcholine in small doses evokes the characteristic post synaptic potential (p s p) (4) the ability of the pre ganglionic fibre to conduct impulses to release acetyl choline and to synthesize acetylcholine *in vitro* disappear on de generation (Banister and Scrase 1950, Hebb and Waites, 1956), while the post synaptic membrane remains responsive to acetyl choline, (5) the ganglion cells cease to respond to acetylcholine if caused to degenerate while the pre ganglionic fibres retain their ability to synthesize and release acetylcholine (6) the post

It has long been known that most if not all of the acetyl choline in nervous tissue is in a difficultly extractable or bound form and can only be liberated by special conditions These conditions have been recently very fully studied by Stone (1955) and are essentially the conditions which are required for the disruption of mitochondria The possibility that acetyl choline *in vivo* is bound to or contained in mitochondrion like particles (Bodian 1942, Feldberg 1945) is interesting in view of the work of Blaschko Hagen and Welsh (1955) on the intra cellular localization of adrenaline in the adrenal medulla Similar considerations apply to choline acetylase the enzyme responsible for the synthesis of acetylcholine (Hebb and Small man, 1956) It is tempting to conclude that acetylcholine syn thesis and storage take place within what Peters (1951) would call the mitochondrial box , thence to be conveyed to the nerve ending to be liberated in response to the appropriate signal

We thus see that transmission at cholinergic nerve endings involves four main processes (1) the synthesis of acetylcholine (2) release of acetylcholine in response to the nerve impulse, (3) combination of acetylcholine with receptor sites and de polarization of the post synaptic membrane (4) destruction of acetylcholine and repolarization of the post synaptic membrane All these processes should be susceptible to biochemical investi gation though obviously with varying degrees of difficulty I shall now attempt to indicate briefly what has so far been achieved under these heads and what remains to be done

Release of Acetylcholine

Process (2) the release of acetylcholine in response to nerve impulses may be dismissed shortly for there is little that can be said at the present time which would not be pure speculation Nevertheless the isolation and characterization of bound acetylcholine offers an experimental approach to this problem Hebb and Whittaker (1958) and Whittaker (1958) have confirmed that most of the acetylcholine in brain tissue is bound to or contained within subcellular particles which have sedi mentation properties grossly similar to those of brain mito- chondria but which are probably specialized particles These

In the sympathetic ganglion, as in other parts of the nervous system and in other tissues, two distinct cholinesterases are to be recognized one characterized by its preferential hydrolysis of acetylcholine at low substrate concentrations which I will refer to as the aceto cholinesterase (Whittaker, 1953b), the other, by its preferential hydrolysis of physiologically active higher acylcholines, which I will refer to, for historical reasons, as the pseudo cholinesterase There is, indeed, nothing 'pseudo' about the action of this enzyme, but it is perhaps too late now to change its name to something more non committal like deutero-cholinesterase Recent studies (Hebb, Silver, Swan and Walsh, 1953, Cavanagh, Thompson and Webster, 1954, Koelle, 1954, Bulbring, Philpot and Bosquanet, 1953) have established that the pseudo cholinesterase is associated with the glial cells If the pre ganglionic fibre is caused to degenerate, the aceto-cholinesterase disappears *pari passu* with the acetylcholine synthesis (Sawyer and Hollingshead, 1945), but the pseudo cholinesterase remains unaffected In the analogous case of the neuromuscular junction, degeneration of the motor nerve reduces, but does not abolish the aceto cholinesterase which warns us not to expect too uniform a pattern at cholinergic synapses throughout the central nervous system

Although acetylcholine synthesis occurs throughout the length of a cholinergic neurone, it is unneces ary to conclude that acetylcholine is also involved in conduction along the fibre Recent studies by del Castillo and Katz (1954, 1956) have shown that miniature end plate potentials occur spontaneously even when transmission is not taking place They have concluded that small quanta' of acetylcholine are continuously being liberated from the motor nerve endings even at rest and that the arrival of a nerve impulse results in an enormous increase in the number of quanta emitted They have drawn attention to recent demonstrations of the presence in several types of nerve endings of sub microscopic 'particles' or vesicles' (cf also Palade, 1954, de Robertis, 1955, de Robertis and Franchi, 1956) and have speculated that the popping or bursting of these could result in a quantized liberation of transmitter substance

It has long been known that most if not all of the acetyl choline in nervous tissue is in a difficultly extractable or bound form and can only be liberated by special conditions These conditions have been recently very fully studied by Stone (1955) and are essentially the conditions which are required for the disruption of mitochondria The possibility that acetyl choline *in vivo* is bound to or contained in mitochondrion like particles (Bodian 1942 Feldberg, 1945) is interesting in view of the work of Blaschko Hagen and Welsh (1955) on the intra cellular localization of adrenaline in the adrenal medulla Similar considerations apply to choline acetylase, the enzyme responsible for the synthesis of acetylcholine (Hebb and Small man, 1956) It is tempting to conclude that acetylcholine synthesis and storage take place within what Peters (1951) would call the mitochondrial 'box , thence to be conveyed to the nerve ending to be liberated in response to the appropriate signal

We thus see that transmission at cholinergic nerve endings involves four main processes (1) the synthesis of acetylcholine, (2) release of acetylcholine in response to the nerve impulse (3) combination of acetylcholine with receptor sites and de polarization of the post synaptic membrane (4) destruction of acetylcholine and repolarization of the post synaptic membrane All these processes should be susceptible to biochemical investi gation though obviously with varying degrees of difficulty I shall now attempt to indicate briefly what has so far been achieved under these heads and what remains to be done

Release of Acetylcholine

Process (2) the release of acetylcholine in response to nerve impulses, may be dismissed shortly for there is little that can be said at the present time which would not be pure speculation Nevertheless the isolation and characterization of bound acetylcholine offers an experimental approach to this problem Hebb and Whittaker (1958) and Whittaker (1958) have confirmed that most of the acetylcholine in brain tissue is bound to or contained within subcellular particles which have sedi mentation properties grossly similar to those of brain mito chondria but which are probably specialized particles These

particles appear to be identical with those containing choline acetylase If they can be isolated in a pure state it should be possible to discover the chemical factors determining the release of acetylcholine from them and perhaps to relate these to the chemical concomitants of electrical activity

Combination of Acetylcholine with Receptor Sites

There is also not much one can say about process (3), the combination of acetylcholine with receptor sites Pharmacologists have been talking about receptor sites for years—one can trace the concept back to Ehrlich The first reaction of the biochemist is to wonder why, if these sites exist, no one has isolated and characterized them Perhaps, however, we are dealing with receptor systems rather than receptor sites—delicately poised metabolic systems which are capable of altering the internal ionic environment of the cell as a result of being stimulated by the sudden availability of the appropriate substrate Nowadays one would look for such a system among the phospholipids, and here it is worth mentioning that phosphatidylserine has recently been reported to be a selective chelating agent for potassium (Solomon, Lionetti and Curran, 1956) In this general connection some recent work by Hokin (1956) may prove to be highly significant

Hokin s studies began with pancreas He measured the incorporation of ^{32}P into various phospholipid fractions during pancreatic secretion *in vitro* He found that agents which specifically stimulate enzyme secretion—including acetylcholine in low concentration—also increase the rate of incorporation of ^{32}P into certain phospholipid fractions This effect has also been detected in other secreting tissues for acetylcholine and adrenaline in salivary gland and for thyrotrophic hormone in thyroid tissue Not all phospholipid fractions are affected equally, the fraction most affected in pancreas and thyroid is the phosphoinositide fraction Similar effects have been found in brain cortex slices and homogenates

Table 1 shows that acetylcholine stimulates the incorporation of ^{32}P into the various phospholipid fractions This effect is small with phosphatidylcholine negligible with phosphatidyl

ethanolamine and phosphatidylserine and greatest with di phosphoinositide It is also marked with the phosphatidic acid fraction, but owing to the small size of this fraction this is again less important than the diphosphoinositide effect

TABLE 1 Effect of acetylcholine on the incorporation of ^{32}P into the phospholipids of guinea pig brain cortex slices (results of Hokin and Hokin 1956)

The system containing in 4 ml Krebs bicarbonate saline 8 mg glucose 80 µc ^{32}P as $NaH_2{}^{32}PO_4$ 400 mg tissue and either 4 mµM atropine sulphate (control) or 240 µM eserine and 40 µM acetylcholine(ACh) was incubated for 3 hr at 39 Counts corrected to specific activity of 10^5 cts/min /µg P for inorganic P in medium

Phospholipid type	% total phospholipid P	Specific activity (cts/min /µg P) Control	ACh
Phosphatidylcholine	47 4	76	126
Phosphatidylethanolamine	27 8	9	12
Phosphatidylserine	19 3	2	3
Diphosphinositide	3 4	710	1 650
Phosphatidic acid	0 7	1 020	2 180
Calcd overall specific activity		70	135

The effects shown in the table were obtained with cortex slices and acetylcholine at 10^{-2} M concentration but other work indicated that the threshold effect was obtained with 10^{-6} to 10^{-5} M acetylcholine and could be demonstrated in particulate preparations as well as slices It thus occurs within the physiological range of concentration of acetylcholine and is evidently an intracellular effect

The significance of these findings is difficult to assess at the present time The most obvious point is that brain cortex is not as far as is known a protein or polypeptide secreting tissue like the other tissues in which this effect has been observed How ever it is known that part of the cell potassium is firmly bound to mitochondria Interference with mitochondrial phospholipid metabolism might well result in changes in cell potassium which could modify the electrical properties of the post synaptic mem brane

SYNTHESIS AND DESTRUCTION OF ACETYLCHOLINE

In contrast to processes (2) and (3), much is now known on the biochemical side regarding processes (1), the synthesis of acetyl choline, and (4), its destruction The main facts are summarized in Figure 2

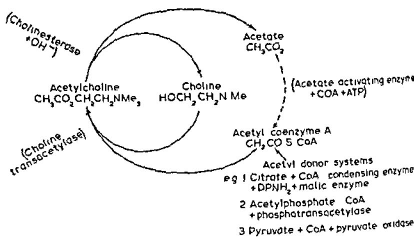


FIG 2 Synthesis and breakdown of acetylcholine

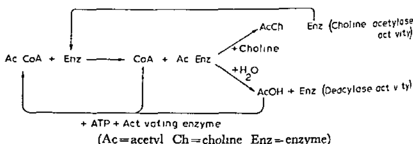
Acetic acid, resulting from the hydrolysis of acetylcholine by cholinesterases, enters the metabolic pool, where it is converted into the acetyl derivative of coenzyme A (CoA), a complex pantothenic acid derivative which is now known to be involved in many different types of acetyla tion reactions (reviewed by Lipmann, 1953) The acetylation of coenzyme A by acetate is cata lysed by the acetate activating enzyme and requires a source of energy in the form of adenosine triphosphate (ATP) Three such activating enzymes are known which between them can activate the whole gamut of fatty acids from acetic to stearic with the formation of the corresponding acyl coenzyme derivatives, and other acids such as benzoic and succinic are also activated in a similar way (Beinert Green, Hele Hift von Korff and Rama krishnan, 1953, Mahler, Wakil and Bock, 1953, Kornberg and Pricer, 1953, Schachter and Taggart, 1953, Kaufman 1953, Sanadi, Gibson and Ayengar 1954) Acetyl coenzyme A is not

only formed from acetate, it is the penultimate product of a number of catabolic processes, including the breakdown of fats and carbohydrates It in turn combines with oxaloacetate to form citrate which is utilized via the tricarboxylic acid cycle In mammalian and avian nervous tissue acetate activation is quantitatively a relatively minor process and there is reason to believe that most of the acetyl coenzyme A comes from other sources I have therefore indicated the connection between acetate and acetyl coenzyme A by a dotted line A considerable proportion of acetate activating enzyme of brain preparations may indeed reside in the supporting tissue since Berry (1955) has shown that acetate activation actually increased in the degenerating proximal portion of sectioned cat sciatic nerves up to the ninety sixth day at which time there had been consider able proliferation of Schwann cells, although acetylcholine syn thesis had disappeared after eight days

The transfer of acetate from acetyl coenzyme A to choline is catalysed by a third enzyme choline acetylase The name cho line acetylase was originally used to denote the complete system which utilized acetate (or a suitable acetate donor like citrate) coenzyme A, ATP and choline for the synthesis of acetyl choline It is now used to denote only the transferring enzyme (Korey de Braganza and Nachmansohn 1951) To emphasize this distinction I shall use the term choline transacetylase rather than choline acetylase when referring to the transferring enzyme alone The separate nature of the activating and transferring enzymes has been amply demonstrated for mammalian brain by Kumagai and Ebashi (1954), who have prepared the latter in a purified form Choline transacetylase is only one of a number of acyl group transferring enzymes now known

The best proof that acetyl coenzyme A is the immediate pre cursor of acetylcholine has come from experiments by Korkes del Campillo, Korey, Stern Nachmansohn and Ochoa (1952) in which squid ganglion choline transacetylase preparations have been coupled with various acetyl coenzyme A generating systems or in which acetyl *coenzyme A in stoichiometric amounts* replaced the donor systems Even in the most highly purified choline transacetylase preparations however, acetyl coenzyme

A is not as efficient a source of acetyl groups for acetylcholine synthesis as is a donor system in which coenzyme A is present in catalytic amounts (Berman, Wilson and Nachmansohn, 1953) Only 30 per cent or less of the acetyl groups of added acetyl coenzyme A are transferred to choline This is probably due to the presence of deacylases (thiolesterases), enzymes catalysing the hydrolysis of acetyl coenzyme A Conceivably, part of this deacylase activity is an inherent property of choline transacetylase itself, just as transferase activity is a property of certain hydrolytic enzymes, notably phosphatases This is equivalent to proposing that water effectively competes with choline for acetyl groups at the enzyme surface as in the following scheme



(Ac=acetyl Ch=choline Enz=enzyme)

If this is correct, the addition of ATP should make acetyl coenzyme A as effective an acetyl donor as ATP plus coenzyme A plus acetate As shown in Table 2 (Whittaker 1953a), ATP does make acetyl coenzyme A more effective as an acyl donor, but it is still not as effective as the acetate plus coenzyme A plus ATP system, suggesting that endogenous acetyl acceptors other than water or choline are involved

Figure 2 also shows some of the acetyl coenzyme A generating systems which have been successfully coupled *in vitro* with choline trans acetylation The condensing enzyme is the enzyme which catalyses the reaction between acetyl coenzyme A and oxaloacetate *to give citrate It is the first of that group* of enzymes which collectively catalyse the events of the tricarboxylic acid cycle In theory it should be possible to reverse this action using citrate as a source of acetyl coenzyme A in practice it was found that this could only be achieved, in coupled systems

if oxaloacetate, the other product of the reaction, were removed by reduction to malate with reduced diphosphopyridine nucleotide ($DPNH_2$) and malic dehydrogenase Other acetyl coenzyme A generating systems which have been used are acetyl phosphate which reacts with coenzyme A to give inorganic P

TABLE 2 Acetyl-coenzyme A as an acetyl donor for acetylcholine synthesis in pigeon brain (*results of Whittaker 1953a*)

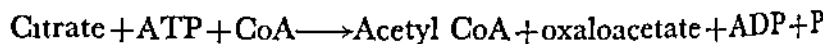
The system contained in 1 16 ml enzyme (resin treated saline extract of 10 mg acetone powder) KCl (50 µM) Na/PO_4 buffer pH 7 0 (20 µM) eserine (1 mg) NaF (12 µM) choline (10 µM) cysteine (25 µM) and acetate (10 µM) ATP (1 5 mg) coenzyme A (0 25 µM) or acetyl coenzyme A (0 25 µM) as indicated
Incubation 37 for 30 min

	Acetylcholine synthesized (µM/g powder/hr)	
Acetate+ATP+CoA	7 50	100
Acetyl CoA	0 62	8
Acetyl CoA+Acetate	0 60	8
Acetyl CoA+ATP	4 90	65
Acetyl CoA+Acetate+ATP	7 50	100

and acetyl coenzyme A in the presence of bacterial phospho transacetylase and pyruvate which is oxidatively decarboxylated to acetyl coenzyme A by the pyruvate oxidation system of heart and other tissues

As I have already indicated there is some doubt that acetate is in fact the main source of acetyl groups for acetylcholine synthesis in the higher vertebrate nervous system From what we know of the metabolism of brain we may assume that pyruvate or ultimately glucose is the main source Acetate is poorly utilized by brain preparations which rapidly oxidize pyruvate and the various cycle intermediates (Peters and Wakelin 1953) and carboxylic acids generally are not considered to pass the blood brain barrier Nevertheless crude choline acetylase preparations from brain can activate acetate for acetylcholine formation they can also utilize citrate Mammalian preparations indeed, utilize citrate more effectively than acetate for this purpose (Balfour and Hebb 1952) though the reverse is

true for invertebrate preparations (Smallman, 1956) One might assume that citrate is converted to acetyl coenzyme A via the condensing enzyme but the available evidence suggests this is not so (Ebashi, 1954) An enzyme has been described in pigeon liver and in brain (Srere and Lipmann, 1953, Balfour, 1955) which catalyses the reaction



$$\text{Citrate} + \text{ATP} + \text{CoA} \longrightarrow \text{Acetyl CoA} + \text{oxaloacetate} + \text{ADP} + \text{P}$$

and another in micro-organisms (Dagley and Dawes, 1953) which catalyses the reaction



$$\text{Citrate} \longrightarrow \text{Acetate} + \text{oxaloacetate}$$

Neither of these enzymes is identical with the condensing enzyme The first might involve citryl coenzyme A as an intermediate, but Srere and Lipmann were unable to establish this

It is quite possible, then, that mammalian brain utilizes citrate as the immediate source of active acetate for acetylcholine synthesis perhaps by means of an enzyme distinct from the condensing enzyme and associated in some special way with the transferring enzyme It is important to realise, however, that the acetate or citrate utilizing systems of crude brain choline acetylase preparations are present in rate limiting concentrations This has been amply demonstrated in work by Hebb (1955) who has obtained rates of acetylcholine synthesis in certain portions of the nervous system three or four times those previously reported by addition of an exogenous source of activating enzyme (a pigeon liver extract) This technique considerably increases the sensitivity with which choline acetylase can be detected in small samples of nervous tissue, and is of particular value in studies of the topographical location of choline transacetylase in the nervous system In this way, it has been possible to demonstrate that certain parts of the central nervous system, e g the dorsal roots are completely devoid of acetylcholine *synthesizing ability—a rather damaging finding* for those who maintain that acetylcholine is involved ubiquitously in nerve conduction (cf Nachmansohn, 1955)

Specificity of Choline Transacetylase

Choline transacetylase is not completely specific for acetyl choline synthesis Its specificity for the choline portion has recently been investigated by Burgen, Burk and Desbarats-Schonbaum (1956) Their results (Table 3) show that consider

TABLE 3 Specificity of rat brain choline acetylase (results of Burgen Burk and Desbarats Schonbaum 1956)

The complete system containing acetylcoenzyme A (AcCoA) (0 75 μM) base (11 μM) TEPP (0 07 μM) EDTA (2 2 μM) K_2PO_4 buffer pH 7 0 (56 μM) and enzyme in 1 ml was incubated at 36 for 30 min Rate of deacetylation of AcCoA measured by fall in optical density at 232 mμ against blank containing no base

$\overset{+}{N}\,C_2H_4OH$			
R	R_1	R_2	Rate (%)
Me	Me	Me	100
Me	Me	Et	143
Me	Et	Et	102
Et	Et	Et	95
H	Me	Me	11
H	H	Me	15
H	H	H	15
Me	Me	n C_3H_7	74
Me	Me	n C_4H_9	67
Me	Me	C_2H_4OH	79
Me	C_2H_4OH	C_2H_4OH	63
C_2H_4OH	C_2H_4OH	C_2H_4OH	16
β methyl	α methylcholine	homocholine	0

able variations in the size of the quaternary nitrogen group are compatible with a high rate of acetylation, indeed the ethyl dimethyl analogue of choline had a higher rate of acetylation than choline itself On the other hand alteration of the hy droxyethyl end as in homocholine and the methylcholines led to a sharp fall in activity The N OH spacing thus seems to be critical

The acyl group specificity of greater physiological interest in view of the occurrence of higher homologues of acetyl coenzyme A in the body and the considerable physiological

true for invertebrate preparations (Smallman, 1956) One might assume that citrate is converted to acetyl coenzyme A via the condensing enzyme but the available evidence suggests this is not so (Ebashi, 1954) An enzyme has been described in pigeon liver and in brain (Srere and Lipmann, 1953, Balfour, 1955) which catalyses the reaction



Citrate + ATP + CoA ⟶ Acetyl CoA + oxaloacetate + ADP + P

and another in micro-organisms (Dagley and Dawes, 1953) which catalyses the reaction



Citrate ⟶ Acetate + oxaloacetate

Neither of these enzymes is identical with the condensing enzyme The first might involve citryl coenzyme A as an intermediate, but Srere and Lipmann were unable to establish this

It is quite possible, then, that mammalian brain utilizes citrate as the immediate source of active acetate for acetylcholine synthesis, perhaps by means of an enzyme distinct from the condensing enzyme and associated in some special way with the *transferring enzyme It is important to realise*, however, that the acetate or citrate utilizing systems of crude brain choline acetylase preparations are present in rate limiting concentrations This has been amply demonstrated in work by Hebb (1955) who has obtained rates of acetylcholine synthesis in certain portions of the nervous system three or four times those previously reported by addition of an exogenous source of activating enzyme (a pigeon liver extract) This technique considerably increases the sensitivity with which choline acetylase can be detected in small samples of nervous tissue, and is of particular value in studies of the topographical location of choline transacetylase in the nervous system In this way, it has been possible to demonstrate that certain parts of the central nervous system e g the dorsal roots are completely devoid of acetylcholine synthesizing ability—a rather damaging finding for those who maintain that acetylcholine is involved ubiquitously in nerve conduction (cf Nachmansohn, 1955)

by this tissue is in reality brought about by three different enzymes (1) the classical choline acetylase, which synthesizes only acetyl and propionylcholine at significant rates (2) an enzyme responsible for the synthesis of the intermediate choline

TABLE 4 Synthesis of *n* acyl cholines by fortified unpurified pigeon brain choline transacetylase

The system contained in 2 7 ml NaCl (200 μM) $MgCl_2$ (20 μM) KCl (100 μM) ATP (10 μM) cysteine (20 μM) coenzyme A (0 1 μM) choline (20 μM) eserine (0 1 μM) fatty acid (20 μM) extract from pigeon liver acetone powder (20 mg) extract from pigeon brain acetone powder (20 mg) and was incubated for 1 hr and pH 7 at 37 Products were separated and identified by chromatography on columns of weak acid ion exchange resin as described by Gardiner and Whittaker (1954) Values for esters other than acetylcholine are exclusive of a small amount of acetyl choline also synthesized

Fatty acid added	Ester synthesized	
	Identity	Quantity (μM/g /hr)
Acetate	acetylcholine	15 0
Propionate	propionylcholine	12 0
Butyrate	butyrylcholine	4 0
Valerate	valerylcholine	3 5
Caproate	caproylcholine	3 0

esters up to caproylcholine (3) the palmitylcholine synthesizing enzyme, which is not involved in the synthesis of any of the lower acyl cholines The situation with respect to the esterifica tion of choline is thus at least as complex as that with respect to choline ester hydrolysis Further knowledge in this direction may well contribute to an understanding of the much discussed role of the pseudo cholinesterases

I suggested some years ago (Whittaker, 1951) that the pseudo cholinesterases, specifically adapted as they are to the hydro lysis of the higher acylcholines may be concerned *in vivo* with the hydrolysis of physiologically active homologues of acetyl choline It seemed likely that some of the higher acyl coenzyme A derivatives formed in the course of the metabolism of the fatty acids should react with choline to form physiologically active acylcholines One such homologue propionylcholine has in deed been identified in ox spleen (Banister Whittaker and

activity possessed by the higher acyl cholines, has also been studied Pigeon brain choline acetylase preparations synthesise propionyl and butyrylcholine if incubated with CoA, ATP and propionate or butyrate in place of acetate (Whittaker, 1953a, Gardiner and Whittaker, 1954) Similar results have been obtained by Nachmansohn and co workers (Korey, de Braganza and Nachmansohn, 1951, Berman, Wilson and Nachmansohn, 1953) with squid ganglion preparations In our own experiments the formation of propionyl and butyrylcholine was checked by chromatographic and biological characterization of the product We felt it was essential to do this because of the ever present possibility that other acids were merely serving as a source of acetate Small amounts of acetylcholine were in fact synthesized by these preparations along with the other esters, these probably arose from endogenous sources of acetate present in the preparations

In these experiments, reliance was placed on the activating enzymes present in the brain preparations Thus the overall specificity was being measured rather than the specificity of the choline transacetylase *per se* Conceivably the limiting factor was the ability of the brain tissue to activate the higher fatty acids Dr Berry and I have now re examined this point using a choline transacetylase preparation fortified with an acyl coenzyme A donor system from pigeon liver capable of activating the homologous series of lower fatty acids Control experiments showed that with the concentration of liver enzyme used the rate of activation was in no case rate determining It will be seen (Table 4) that choline esters of all the *n* fatty acids up to *n* caproate, the highest acid tested, were synthesized, though, as with the unfortified preparations, at rates decreasing with increasing chain length Pigeon liver preparations alone were unable to synthesize choline esters from these shorter chain fatty acids However both pigeon liver and brain were able to synthesize palmitylcholine from added palmitic acid and choline A compound with properties similar to palmitylcholine has recently been discovered by Kennedy (1956) in the phospholipid fraction of brain and liver Purification of pigeon brain by the method of Ebashi (1954) revealed that acyl choline synthesis

the isolated head was noted The ability of the drugs to cause aceto cholinesterase *inhibition in vivo* was *ingeniously tested* by measuring the threshold dose for potentiating the twitch height obtained by stimulating the tibialis muscle via the motor nerve This potentiation of twitch height is a phenomenon which

TABLE 5 The effect of selective cholinesterase inhibitors on arousal reaction and twitch height in the cat (results of Desmedt and LaGrutta 1955)

	Mean doses (μg) for	
	Activation (arousal) of sleeping cortex	Potentiation of indirect twitch of tibialis
Ro 2-0683	5	95
DFP	12	205
Ro 2 1250	70	12
BW 284C51	200	5

depends on inactivation of the aceto cholinesterase at the neuromuscular junction The selective nature of the inhibitors for the two cholinesterases of cat s brain was also checked by conventional Warburg methods Results were obtained with two selective inhibitors of each sort one of each pair being selected for its predominantly water soluble character, the other for its lipoid solubility In this way the factor of differential solubility or penetrability was controlled It will be seen from the results presented in Table 5 that the two pseudo cholinesterase inhibitors were effective in eliciting the arousal reaction in extremely low doses, whereas much higher doses of the aceto cholinesterase inhibitors were needed The effect of the pseudo-cholinesterase inhibitors could not have been due to the inhibition of aceto cholinesterase since much higher doses of these same inhibitors were needed to inhibit aceto-cholinesterase as measured by *the potentiation of twitch height The aceto* cholinesterase inhibitors were probably acting by way of inhibition of pseudo-cholinesterase since the doses of these inhibitors required to elicit the arousal reaction were much higher than those required to inhibit aceto-cholinesterase as measured by the potentiation of twitch height

Wijesundera, 1953, Gardiner and Whittaker, 1954) However, in a rather extensive study of other tissues of ox, sheep, goat and rabbit, we have been unable to identify any physiologically active ester of choline other than acetylcholine (Keyl and Whittaker, 1954) This suggests that the accumulation of these higher homologues is effectively suppressed by the presence of the pseudo cholinesterases Why this does not happen in the case of ox spleen is not known We have attempted to test this hypothesis by repeatedly injecting propionate into rabbits chronically intoxicated with DFP so as to maintain blood levels of propionate of about 2 mg /100 ml over a period of two to three weeks We hoped that the inhibition of pseudo-cholin esterase would allow propionylcholine to accumulate However, no propionylcholine could be demonstrated at the end of this period in brain or intestine, the two tissues containing the highest acetylcholine equivalences, in spite of a considerable rise in the acetylcholine equivalence of the intestine Such experiments suffer from the technical limitation that it is extremely difficult to bring about a significant degree of depres sion of tissue pseudo cholinesterase with the currently available selective inhibitors without at the same time depressing the aceto cholinesterase level to danger point Some of the best selective inhibitors *in vitro* also have undesirable side reactions *in vivo* The plasma cholinesterase of our rabbits never fell below 25 per cent of controls

Whether or not the function of the pseudocholinesterases is as I have suggested evidence continues to accumulate that pseudo cholinesterases play some part in nervous activity A recent example is Desmedt and LaGrutta s (1955) finding that selective inhibitors of pseudo-cholinesterases in low concentra tion can evoke the cortical potentials in sleeping cats which are characteristic of the waking state This arousal' reaction is produced by selective inhibitors of aceto cholinesterase only in much higher concentrations where they are probably no longer selective The technique is briefly as follows The drugs were in jected into the external carotid artery of unanaesthetized cats with a spinal transection at C1 The minimum dose required to accelerate and desynchronize the spontaneous rhythms of

the isolated head was noted The ability of the drugs to cause aceto cholinesterase inhibition *in vivo* was ingeniously tested by measuring the threshold dose for potentiating the twitch height obtained by stimulating the tibialis muscle via the motor nerve This potentiation of twitch height is a phenomenon which

TABLE 5 The effect of selective cholinesterase inhibitors on arousal reaction and twitch height in the cat (results of Desmedt and LaGrutta 1955)

	Mean doses (μg) for	
	Activation (arousal) of sleeping cortex	Potentiation of indirect twitch of tibialis
Ro 2-0683	5	95
DFP	12	205
Ro 2 1250	70	12
BW 284C51	200	5

depends on inactivation of the aceto cholinesterase at the neuromuscular junction The selective nature of the inhibitors for the two cholinesterases of cat s brain was also checked by conventional Warburg methods Results were obtained with two selective inhibitors of each sort, one of each pair being selected for its predominantly water soluble character the other for its lipoid solubility In this way the factor of differential solu bility or penetrability was controlled It will be seen from the results presented in Table 5 that the two pseudo-cholinesterase inhibitors were effective in eliciting the arousal reaction in extremely low doses, whereas much higher doses of the aceto cholinesterase inhibitors were needed The effect of the pseudo-cholinesterase inhibitors could not have been due to the inhibi tion of aceto cholinesterase since much higher doses of these same inhibitors were needed to inhibit aceto-cholinesterase as measured by the potentiation of twitch height The aceto-cholinesterase inhibitors were probably acting by way of inhi bition of pseudo-cholinesterase since the doses of these inhibi tors required to elicit the arousal reaction were much higher than those required to inhibit aceto-cholinesterase as measured by the potentiation of twitch height

The activation of brain potentials by way of pseudo-cholin esterase inhibition is especially interesting in view of the location of brain pseudo cholinesterase in the neuroglia and the possibility that the supporting cells of the central nervous system have a considerable fatty acid metabolism of their own which may be connected with myelin formation Conceivably the pseudo cholinesterase of the central nervous system is concerned with the suppression of the formation of physiologically active higher acyl cholines which might otherwise interfere with acetylcholine function in neighbouring nerve cells Possibly significant in this connection is the ability of pro pionylcholine and butyrylcholine to awake the sleeping cat's brain in low concentrations

To these considerations must now be added the further possi bility that choline esters of higher fatty acids may make a significant contribution to the composition of the brain lipids and that their synthesis and breakdown are brought about by mechanisms closely paralleling the acetylcholine system of cholinergic nerve cells The well known demyelinating action of certain anti cholinesterases might usefully be reinvestigated along these lines, and the possibility that the two systems may interact at the level of the post synaptic membrane suggests intriguing vistas for the future

It has been known for some time that the three components of the acetylcholine system—choline acetylase *aceto-cholin* esterase and acetylcholine itself—are distributed in a patchy manner in the central nervous system Unfortunately these three entities have never been quantitated by one investigator at the same time in identical tissue samples therefore less satisfactory comparisons between the results of different workers have had to be made Nevertheless, these comparisons show a satisfying degree of correlation It is easiest to see this if individual tracts are followed up As an illustration of the degree of correlation obtained I have plotted in Figure 3 the distribution of the three components for the somatic sensory and pyramidal (motor) tracts as tabulated by Burgen and Chipman (1951) from their own results for cholinesterase and those of Feldberg and Vogt (1948), and MacIntosh (1941), for choline acetylase and acetyl

choline respectively It will be seen that regions of high aceto cholinesterase activity (white blocks) are also areas of high choline acetylase activity (hatched blocks) and high acetylcholine content (dotted blocks) and vice versa Similar correlations have been found for the optic, acoustic and olfactory tracts and for

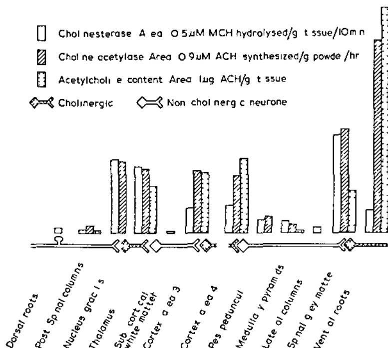


Fig 3 Distribution of acetylcholine system along sensory and motor tracts of the dog MCH = acetyl β methylcholine ACH = acetylcholine

portions of the extra pyramidal system More accurate absolute values for the choline transacetylase distribution have recently been obtained by Hebb and Silver (1956) using the technique of fortified preparations already described without, however, altering the fundamental picture

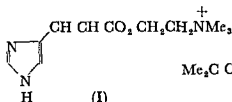
These results strongly suggest that some but not all of the neurones of the central nervous system are cholinergic The cor relation between aceto-cholinesterase and choline acetylase is certainly not perfect but this can be accounted for by assuming

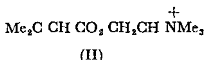
that cholinergic neurones of the central nervous system, like sympathetic pre ganglionic fibres and the somatic motor nerves, synthesize acetylcholine throughout their length, while cholin esterase activity may be concentrated at their endings If this is so, the distribution of choline acetylase, especially, gives us a clue as to which neurones are cholinergic, and the result is obtained that in many tracts there appears to be an alternation between cholinergic and non cholinergic neurones This has been indicated diagrammatically in Figure 3 There are also interesting phylogenetic differences which deserve to be more fully explored For example, all parts of the pigeon brain—hemispheres, optic lobes and cerebellum—are very rich in aceto-cholinesterase (Whittaker, 1953b) It is well known that the avian fore brain—derived as it is from the archistriatum of the primitive brain—is homologous with the mammalian basal ganglia also regions particularly rich in aceto cholinesterase Pallial tissue, represented by the less active mammalian cortex, is but poorly represented in the avian fore brain The optic lobes represent an enormous development of the region of the primitive brain which in mammals is represented by the cor pora quadrigemina, thalamus and geniculate bodies, again areas rich in cholinesterase In the cerebral ganglia of inverte brates—insects, crustaceans and cephalopods—extremely high concentrations of the components of the acetylcholine system are encountered It is possible that cholinergic transmission is a primitive kind of transmission which has been rather fully utilized in the avian brain, but which has been in part super seded in the mammals

Other Active Esters of Choline

Although the emphasis of much recent work on acetylcholine has been on its role in the central nervous system we should not lose sight of the fact that acetylcholine is an extremely widely distributed substance in nature and that it is by no means the only physiologically active ester of choline to occur there We can now add two new ones to the list (Keyl, Michaelson and Whittaker, 1957) One of these occurs in the hypobranchial gland of the predatory marine gastropod, *Thais floridana*, the

other in the same organ of the common whelk of our English coasts, *Buccinum undatum* Neither is identical with urocanyl choline (I), identified by Erspamer and Benati (1953 1954) in the Mediterranean whelks *Murex trunculus*, *M brandaris* and *Tritonalia erinacea* by ourselves in other, American North Atlantic *Muricidae* (Whittaker and Michaelson, 1954) and given clinical application by de Blasi and Leone (1955)





The ester from *Thais floridana* has been identified as 3 3 dimethylacrylylcholine (II) on the basis of gas chromatography of the volatile fatty acid liberated from it by hydrolysis, and other tests (Whittaker, 1957) It is interesting to speculate that histidine and valine respectively may be the precursors in the biosynthesis of these esters this is supported by our observation that *T lapillus* contains a powerful α deaminase Oddly enough, none of the *Muricidae* so far examined has discovered the advantage of reducing the side chain of urocanylcholine to give imidazole propionylcholine, an ester 30 times more active on the frog rectus abdominis muscle than urocanylcholine Perhaps when we know more about the function of these esters in the snail and can evaluate their relative activitie on a test organ more physiological—from the snail s point of view—than the frog rectus, these interrelationships will become more intelligible

However that may be the fuller investigation of these naturally occurring choline esters their function and metabolism, may be expected to provide some interesting analogies with the acetylcholine system with which this review has been mainly concerned

REFERENCES

BALFOUR W E (1955) *J Physiol* **129** 81P

BALFOUR W E and HEBB C (1952) *J Physiol* **118** 94

BANISTER J and SCRASE M (1950) *J Physiol* **111** 347

BANISTER J WHITTAKER V P and WIJESUNDERA S (1953) *J Physiol* **121** 55

BEINERT H GREEN D E HELE P HIFT H H VON KORFF R W and RAMAKRISHNAN C V (1953) *J biol Chem* **203** 35

BERMAN R WILSON I B and NACHMANSOHN D (1953) *Biochim biophys Acta* **12** 315

BERRY J F (1955) Personal communication

BLASCHKO H HAGEN P and WELSH A D (1955) *J Physiol* **129** 27

BODIAN D (1942) *Physiol Rev* **22** 146

BULBRING, E PHILPOT E J and BOSQUANET F D (1953) *Lancet* **264** 865

BURGEN A S V, BURK G and DESBARATS SCHONBAUM (1956) *Brit J Pharmacol* **11** 308

BURGEN A S V and CHIPMAN L M (1951) *J Physiol* **114** 296

CAVANAGH J B THOMPSON R H S and WEBSTER G R (1954) *Quart J exp Physiol* **39** 185

DAGLEY S and DAWES E A (1953) *Nature Lond* **172** 345

DE BLASI S and LEONE U (1955) *Minerva anestesiol* **21** 137

DEL CASTILLO J and KATZ B (1954) *J Physiol* **124** 553 560

DEL CASTILLO J and KATZ B (1956) *Progr Biophys* **6** 121

DE ROBERTIS E (1955) *Anat Rec* **121** 284

DE ROBERTIS E and FRANCHI C M (1956) *J biochem biophys Cyt* **2** 307

DESMEDT J E and LAGRUTTA G (1955) *J Physiol* **129** 46P

EBASHI S (1954) *Jap J Pharmacol* **4** 32

ERSPAMER V and BENATI O (1953) *Science* **117** 161

ERSPAMER V and BENATI O (1954) *Biochem Z* **324** 66

FELDBERG W (1945) *Physiol Rev* **25** 596

FELDBERG W and VOGT M (1948) *J Physiol* **107** 372

GARDINER J E and WHITTAKER V P (1954) *Biochem J* **58** 24

HEBB C O (1955) *Quart J exp Physiol* **40** 176

HEBB C O and SILVER A (1956) Personal communication

HEBB C O SILVER A. SWAN A A B and WALSH F G (1953) *Quart J exp Physiol* **38** 185

HEBB C O and SMALLMAN B N (1956) *J Physiol* **134** 385

HEBB C O and WAITES G M H (1956) *J Physiol* **132** 667

HEBB C O and WHITTAKER V P (1958) *J Physiol* In press

HOKIN L E (1956) *Can J Biochem Biophys* **34** 349

HOKIN L E and HOKIN M R (1956) *Biochim biophys Acta* **18** 107

KAUFMAN S (1953) *Fed Proc* **12** 704

KENNEDY E P (1956) *Can J Biochem Physiol* **34** 334

KEYL J MICHAELSON I A and WHITTAKER V P (1957) *J Physiol* In press

KEYL J and WHITTAKER V P (1954) Unpublished observations
KOELLE G B (1954) *J comp Neurol* **100**, 211
KOREY S R DE BRAGANZA B and NACHMANSOHN D (1951) *J biol Chem* **189** 705
KORKES S DEL CAMPILLO A KOREY S R STERN J R NACHMANSOHN D and OCHOA S (1952) *J biol Chem* **198** 215
KORNBERG A and PRICER W E (1953) *J biol Chem* **204** 329 345
KUMAGAI H and EBASHI S (1954) *Nature Lond* **173** 871
LIPMANN F (1953) *Fed Proc* **12** 673
MACINTOSH F C (1941) *J Physiol* **99** 436
MAHLER H R WAKIL S J and BOCK R M (1953) *J biol Chem* **204** 453
NACHMANSOHN D (1955) In Fulton J F *Textbook of Physiology* 17th edn W B Saunders Co Philadelphia and London
OGSTON A G (1952) Quoted by Eccles J C *The Neurophysiological Basis of Mind* Clarendon Press Oxford
PALADE G E (1954) *Anat Rec* **118** 335
PETERS R A (1951) *Proc roy Soc B* **139** 143
PETERS R A and WAKELIN R W (1953) *J Physiol* **119** 4 1
SANADI D R GIBSON D M and AYENGAR P (1954) *Biochim biophys Acta* **14** 434
SAWYER G H and HOLLINGSHEAD W H (1945) *J Neurophysiol* **8** 137
SCHACHTER D and TAGGART J V (1953) *J biol Chem* **203** 925
SMALLMAN B N (1956) *J Physiol* **132** 343
SOLOMON A K LIONETTI F and CURRAN P F (1956) *Nature Lond* **178** 582
SRERE P A and LIPMANN F (1953) *J Amer chem Soc* **75** 4847
STONE W E (1955) *Arch Biochem Biophys* **54** 181
THESLEFF S (1955) *Nature Lond* **175** 594
WHITTAKER V P (1951) *Physiol Rev* **31** 312
WHITTAKER V P (1953a) Unpublished observations
WHITTAKER V P (1953b) *Biochem J* **54** 660
WHITTAKER V P (1957) *Biochem J* **66** 35P
WHITTAKER V P and MICHAELSON I A (1954) *Biol Bull* **107** 304
WHITTAKER V P (1958) *Biochem J* **68** 21P

XII

Factors Influencing the Action of Neuromuscular Blocking Substances

ELEANOR ZAIMIS

NEUROMUSCULAR blocking drugs have proved to be clinically useful and of great interest to pharmacologists and physiologists, tubocurarine gallamine, decamethonium and suxamethonium holding the centre of the picture All *these are quaternary ammonium compounds and their pharma*cological actions both in animals and man have been studied in great detail (Paton and Zaimis 1952, Zaimis 1954, Bovet and Bovet Nitti, 1955, Goodman and Gilman, 1955, Robson and Keele 1956)

For this reason I will restrict myself to the discussion of some factors which influence their action and which may be of importance in their clinical use

Species Differences

Tubocurarine and gallamine interrupt neuromuscular transmission in all mammals by competition with acetylcholine They do not interfere with the release of acetylcholine from the motor nerve endings and have no effect on the electrical properties of the end plate or muscle fibre These compounds apparently combine reversibly with the end plate receptors and therefore block the normal points of attachment of acetylcholine

The mode of action of decamethonium and suxamethonium on the other hand is not uniform In some mammalian species such as man and cat these compounds interrupt neuromuscular transmission by 'mimicking acetylcholine as was first realized

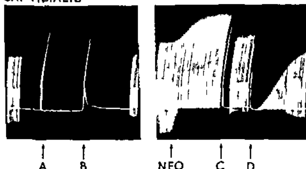


FIG 1 In this and the following figures the tracings are simultaneous recordings of maximal twitches in response to indirect stimulation elicited once every 10 seconds

A *and* B—*contractions in response to acetylcholine injected* intra arterially C and D—neuromuscular block produced by acetylcholine in the presence of neostigmine NEO—neostigmine methylsulphate

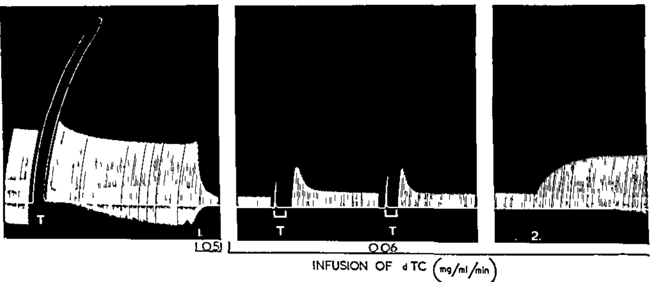


FIG 2 Cat tibialis T—tetanus 50 shocks per second for 15 seconds every ten minutes recorded on a faster drum At 1 a slow continuous infusion of tubocurarine was started At 2 250 μg of neostigmine methylsulphate injected intravenously

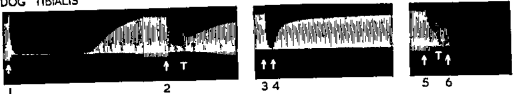


Fig 3 At 1 2 and 4 130 µg of decamethonium diiodide At 3 250 µg neostigmine methylsulphate At 5 0 5 mg of tubocurarine chloride T—tetanus 50 shocks per second

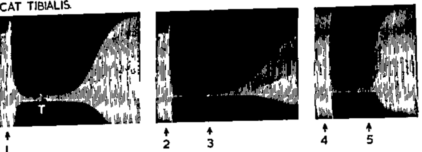


Fig 4 At 1 2 3 and 5 1 mg decamethonium diiodide At 4 0 5 mg of neostigmine methylsulphate At 6 300 µg of tubocurarine chloride T—tetanus 50 shocks per second

PLATE XVIII

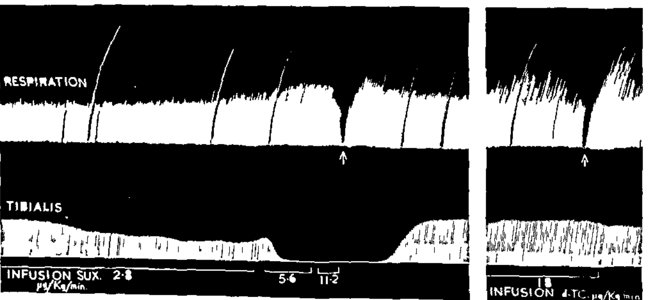


FIG 5 Dog At arrows 0 5 mg neostigmine methylsulphate

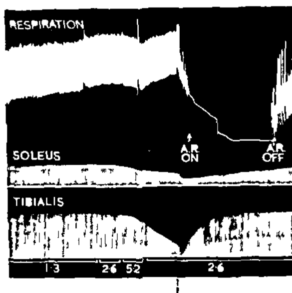


FIG 6 Cat Continuous infusion of suxamethonium diiodide At ↑ 1 mg of tubocurarine chloride A R —artificial respiration

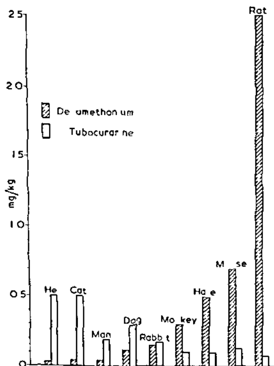


FIG 7 Relative potency of decamethonium and tubocurarine in different species The dose indicated in mg /kg is that required to produce a 90–100 per cent ne omuscular block in all species but mouse where it represents the ED_{50} using the righting reflex test

PLATE XX

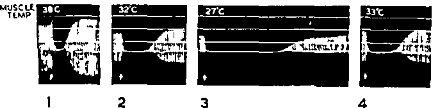


Fig 8 Cat At arrows 1 3 and 4 150 μg of suxamethonium diiodide was administered intravenously



Fig 9 Cat At arrows 1 2 and 3 120 μg decamethonium diiodide was administered intravenously

PLATE XXI

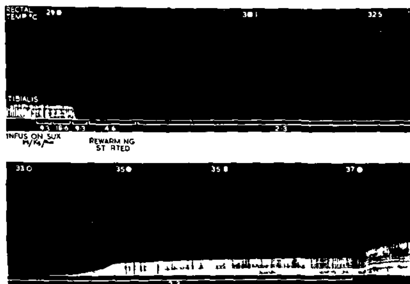


FIG 10 Cat tibialis Continuous infusion of suxamethonium diiodide (sux)

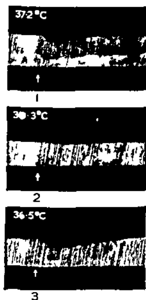


Fig. 11. Cat tibialis. At arrows 1, 2 and 3, 0.5 mg tubocurarine chloride was administered intravenously.

during the study of the mode of action of decamethonium (Zaimis, 1951) During activity the release of acetylcholine from the motor nerve endings its action on the motor end plate and its destruction by the enzyme is a very rapid process completed in a few milliseconds When however very large amounts of acetylcholine accumulate, depolarization of the motor end plates persists during which electrical inexcitability develops and neuromuscular block sets in Consequently acetyl choline itself may (*a*) initiate contraction or (*b*) interrupt trans mission Plate XV, Figure 1 shows these two effects An arterial injection of acetylcholine causes a quick contraction of the tibialis muscle but, in the presence of neostigmine, its action persists and neuromuscular block results

Burns and Paton (1951), using the gracilis muscle in the intact cat, demonstrated that this likeness to acetylcholine rests in the ability of decamethonium to cause a persistent depolariza tion of the end plate region Using the same technique, they showed that all the principal features of block by decame thonium can be reproduced with acetylcholine in the presence of an anticholinesterase drug

In other mammalian species however, including the monkey, dog rabbit and hare, decamethonium and suxamethonium ex hibit initially an acetylcholine like action but during the block ing process their action changes into that of a substance com peting with acetylcholine When first observed this mode of action was described as dual (Zaimis, 1953)

Neuromuscular transmission may thus be interrupted by radically different modes of action Any one of these compounds if injected intravenously into an unanaesthetized animal will produce a flaccid paralysis superficially indistinguishable from that produced by any other But if the animals are anaesthe tized a nerve muscle preparation set up, the indirectly stimu lated contractions of their muscles recorded and the same drugs administered the characteristics of the neuromuscular block they produce are different

The following are the more relevant characteristics of a sub stance blocking by competition with acetylcholine First, the muscle shows an increased sensitivity to subsequent doses

Second, a tetanus produced while the muscle is under the influence of such a drug is not well sustained It gives rise to a transient contraction of the muscle which yields rapidly to complete neuromuscular block for the remaining duration of the tetanus On returning to single twitches a striking and long lasting restoration of the transmission is seen Third, the block is readily antagonized by *neostigmine* or any other substance which sensitizes the motor end plate to acetylcholine or inhibits the enzyme and allows acetylcholine to accumulate (Plate XVI, Figure 2)

When substances interrupt neuromuscular transmission by mimicking acetylcholine the paralysis they produce is usually preceded by potentiation of the maximal twitch and by spontaneous fasciculations A tetanus, if present during the block, is well sustained and does not antagonize it Neostigmine has little effect upon the course of the block or may deepen and prolong it In contrast the block is effectively antagonized by tubocurarine and finally the muscles show the same sensitivity to at least the three or four subsequent doses Plate XVII, Figure 3 shows all these characteristics

I must admit that there are not many mammalian species in whose muscles suxamethonium and decamethonium exhibit a pure 'acetylcholine like action In fact, the only ones so far observed are the cat and man But in such muscles the characteristics of the blockade remain unchanged whether large, small or repeated doses are given or whether the muscle is kept paralysed for several hours

Whereas skeletal muscles when blocked by depolarization, are very sensitive to decamethonium and suxamethonium, they are relatively insensitive when the same compounds interrupt neuromuscular transmission by a dual mode of action Moreover, not only are they relatively insensitive to the first dose but they exhibit a rapidly decreasing sensitivity to repeated doses An analysis of the characteristics of the block shows that it has two components one due to the initial acetylcholine like action, the other to a curare like action The block is preceded by potentiation of the maximal twitch a feature peculiar to a substance capable of depolarizing the motor end plates But a

tetanus produced during the block is not well sustained and antagonizes it The block is deepened by tubocurarine and is readily antagonized by neostigmine In other words one is confronted with a type of block which has some of the characteristics of a depolarization block and some of the characteristics of a competitive one But the striking decreasing sensitivity of the muscles to subsequent doses is a totally new feature absent in either of the two basic mechanisms when they appear in their pure form (for details see Zaimis, 1953) Plate XVII Figure 4 shows the characteristics of the dual mode of action of decamethonium and Plate XVIII Figure 5 demonstrates the curare like effect of suxamethonium both in the dog In the first part of the experiment illustrated by Figure 5, suxamethonium was administered by a slow infusion until it produced respiratory arrest and a 100 per cent block of the tibialis muscle 0 5 mg of neostigmine methyl sulphate administered intravenously readily antagonized the action of suxamethonium in both tibialis and respiratory muscles Three hours later tubocurarine was administered and as was expected neostigmine produced the same antagonism

These results demonstrate that a single substance may produce different types of neuromuscular block in various mammalian species It follows that differences must exist between the muscle membranes of these species in spite of the apparent similarity of their reaction to acetylcholine But where these differences reside remains a matter of speculation

Differences in Various Muscles of the same Species

Further work has demonstrated that differences between neuromuscular junctions exist not only between the muscles of different species but even between various muscles of the same species For example in the cat decamethonium and suxamethonium block the tibialis by an acetylcholine like action but they block soleus by a dual mode of action (Jewell and Zaimis 1954)

Plate XIX Figure 6 gives an idea of the complicated situation which arises when such differences exist in the various muscles of the same species In this experiment the respiration of the

animal and the contractions of the tibialis and soleus muscles were recorded simultaneously, while the rate of suxamethonium infusion was adjusted to produce an 80 per cent block in tibialis At this concentration of suxamethonium the respiratory muscles were little affected and soleus muscle blocked only by 50 per cent One mg of tubocurarine, administered intravenously, readily antagonized the block in tibialis but deepened that of soleus and paralysed the respiratory muscles It is quite obvious that the actions of tubocurarine and suxamethonium are antagonistic in the tibialis muscle of this animal, but both exhibit the same mode of action in soleus Respiratory muscles in the cat are known to be relatively resistant to depolarizing drugs (Paton and Zaimis, 1951), possibly because, as in the case of soleus, these drugs exhibit a dual mode of action

The phases of depolarization and competition in muscles in which suxamethonium and decamethonium exhibit a 'dual' mode of action are not always the same In some muscles the depolarization phase predominates, in others the competing one Nor is the dual mode of action which decamethonium and suxamethonium exhibit the same The balance between the depolarizing and competitive propensities with suxamethonium appears weighted in favour of the former, whereas with decamethonium it is the competitive tendency which appears dominant In other words, of the two, it is suxamethonium which mimics acetylcholine more faithfully

One thing should be clearly understood, that the mode of action of suxamethonium and decamethonium *never* changes during the course of a single experiment If the species or the individual muscles under study respond to these compounds by pure depolarization the response will remain unaltered throughout the experiment In nine years of work I have never seen a muscle which was blocked by a pure acetylcholine like action at first and later showed a dual' block

Interpretation of Results

The results described provide an explanation for the varying sensitivity of different species to decamethonium and suxamethonium, in contrast to their relatively uniform reaction to

tubocurarine (Plate XIX, Figure 7) Sensitivity is at its highest where the compounds interrupt neuromuscular transmission by an acetylcholine like action Immediately the dual' mode of action appears, the muscles become more resistant to the same drugs most probably because these two modes of action are antagonistic

Thus it is fairly clear that the ability of a compound to pro duce a neuromuscular block by several modes of action, although a very interesting phenomenon, is a complicating factor in the handling of such drugs Thesleff (1955a b) from electrophysiological experiments on the isolated diaphragm of the rat and the sartorius muscle of the frog has questioned the existence of a block due solely to depolarization While not denying the similarity of decamethonium and suxamethonium to acetylcholine, he claims that depolarization is not sustained during the whole period of the block But we know that the results obtained with these compounds in one species are not necessarily valid for another, and furthermore that there are species in which the results obtained in one muscle are not repeated in all muscles Thesleff in his papers disagrees with the results of Burns and Paton (1951) but although his findings and interpretations may be correct for the isolated muscles of the frog and the rat it seems doubtful whether they can safely be transferred to the intact cat

Another question I would like to raise is that of experiments on isolated muscles Here the difficulty at once arises that there is no means of proving that the physiological properties of isolated tissues suffer no alteration through being separated from the body of which these tissues are an integral part Further in order to study the effects of different drugs on such preparations the substances are added to the fluid surrounding the isolated tissue It is obvious that there can be no real resemblance between such an application and the contact which drugs make with the motor end plates when carried by the circulation Thus species and muscle differences the differ ent experimental techniques used etc have to be considered very carefully before the results obtained are interpreted and especially before they are transferred to man The answer will

animal and the contractions of the tibialis and soleus muscles were recorded simultaneously, while the rate of suxamethonium infusion was adjusted to produce an 80 per cent block in tibialis At this concentration of suxamethonium the respiratory muscles were little affected and soleus muscle blocked only by 50 per cent One mg of tubocurarine administered intravenously, readily antagonized the block in tibialis but deepened that of soleus and paralysed the respiratory muscles It is quite obvious that the actions of tubocurarine and suxamethonium are antagonistic in the tibialis muscle of this animal, but both exhibit the same mode of action in soleus Respiratory muscles in the cat are known to be relatively resistant to depolarizing drugs (Paton and Zaimis, 1951), possibly because, as in the case of soleus, these drugs exhibit a 'dual' mode of action

The phases of depolarization and competition in muscles in which suxamethonium and decamethonium exhibit a dual mode of action are not always the same In some muscles the depolarization phase predominates, in others the competing one Nor is the dual mode of action which decamethonium and suxamethonium exhibit the same The balance between the depolarizing and competitive propensities with suxamethonium appears weighted in favour of the former, whereas with decamethonium it is the competitive tendency which appears dominant In other words of the two, it is suxamethonium which mimics acetylcholine more faithfully

One thing should be clearly understood, that the mode of action of suxamethonium and decamethonium *never* changes during the course of a single experiment If the species or the individual muscles under study respond to these compounds by pure depolarization, the response will remain unaltered throughout the experiment In nine years of work I have never seen a muscle which was blocked by a pure acetylcholine like action at first and later showed a dual block

Interpretation of Results

The results described provide an explanation for the varying sensitivity of different species to decamethonium and suxamethonium, in contrast to their relatively uniform reaction to

prolongs the action of depolarizing drugs, *the nature of the blockade* is in no way affected however long the paralysis lasts

Cooling has less effect on the action of substances blocking by competition with acetylcholine In fact when small doses of such drugs are used producing less than a 40 per cent reduction of the maximal twitch tension the degree of block is actually smaller at lower temperature Plate XXII, Figure 11 illustrates the results obtained with a small dose of tubocurarine Holmes, Jenden and Taylor (1951), in experiments carried out on the isolated diaphragm of the rat have already demonstrated that cooling reduces the effect of tubocurarine When tubocurarine is administered in larger doses, the magnitude of the effect is not significantly changed by cooling, but the *duration* of the blockade may be prolonged But one must remember that tubocurarine a substance only slowly destroyed and excreted (Marsh, 1952), produces progressively larger responses even at normal body temperatures

An attempt was made to express qualitatively the effect of cooling on the action of the two groups of drugs by measuring the increase in 50 per cent recovery time produced by a change in temperature from approximately 36 C to 31 C It was found that with larger doses of tubocurarine a fall in temperature of this order approximately doubled the 50 per cent recovery time However, if the temperature was raised again the effect was only slightly reversed On the other hand with decamethonium and suxamethonium a similar fall in temperature caused a fourfold increase in 50 per cent recovery time and this effect was completely reversed on rewarming

From these results it is clear that cooling has a pronounced effect on the action of depolarizing drugs The fact that a block produced by decamethonium is more or at least as much, prolonged as one produced by suxamethonium suggests that the prolongation of the effect cannot be due to inactivation of enzymes at the lower temperatures for if this were so the suxamethonium block should be the most affected This result together with the fact that substances interrupting the neuromuscular transmission by competition with acetylcholine are less affected by cooling suggests that cooling influences the

come only when the same skilled electrophysiological techniques which scientists use now on isolated muscles of frogs and rats can be applied to the muscles of the intact man or cat

THE EFFECT OF MUSCLE TEMPERATURE ON THE ACTION OF NEUROMUSCULAR BLOCKING DRUGS

Work done recently by my colleagues and myself has brought to light a new factor which underlines once more the differences between the modes of action of these drugs (Bigland, Goetzee, Maclagan and Zaimis, 1958)

The effect of body—and particularly muscle—temperature on the action of neuromuscular blocking drugs was studied on animals whose body temperature was varied at the rate of 4°C per hour by passing hot or cold water through a thin rubber bag inserted into the abdominal cavity or by modifying the external heating This method made it possible to produce a slow but steady fall in body temperature without perceptible shivering

When the neuromuscular blocking effects of successive doses of suxamethonium and decamethonium were tested at different temperatures it was found that the block produced by the same dose was greater both in *magnitude* and *duration* the lower the temperature Plate XX, Figure 8 illustrates the results obtained in an experiment in which suxamethonium was used The first dose administered when the muscle temperature was 38 C, produced the usual short lasting interruption of neuromuscular transmission As the muscle temperature fell, the duration of the blockade became more and more prolonged Such effects were regularly obtained and were completely reversible on rewarming the animal Plate XX, Figure 9 illustrates similar results with decamethonium

Plate XXI Figure 10 illustrates the results of an experiment in which suxamethonium was administered by a slow infusion after the body temperature had been lowered to 29°C The rate of infusion was adjusted to produce a 100 per cent block and kept constant while the animal was rewarmed It can be clearly seen that the magnitude of the neuromuscular block followed the temperature variations faithfully Although cooling considerably

prolongs the action of depolarizing drugs, *the nature of the blockade* is in no way affected, however long the paralysis lasts

Cooling has less effect on the action of substances blocking by competition with acetylcholine In fact, when small doses of such drugs are used producing less than a 40 per cent reduction of the maximal twitch tension the degree of block is actually smaller at lower temperature Plate XXII Figure 11 illustrates the results obtained with a small dose of tubocurarine Holmes Jenden and Taylor (1951) in experiments carried out on the isolated diaphragm of the rat, have already demonstrated that cooling reduces the effect of tubocurarine When tubocurarine is administered in larger doses the magnitude of the effect is not significantly changed by cooling but the *duration* of the blockade may be prolonged But one must remember that tubocurarine, a substance only slowly destroyed and excreted (Marsh 1952) produces progressively larger responses even at normal body temperatures

An attempt was made to express qualitatively the effect of cooling on the action of the two groups of drugs by measuring the increase in 50 per cent recovery time produced by a change in temperature from approximately 36 C to 31 C It was found that with larger doses of tubocurarine a fall in temperature of this order approximately doubled the 50 per cent recovery time However if the temperature was raised again the effect was only slightly reversed On the other hand with decamethonium and suxamethonium a similar fall in temperature caused a fourfold increase in 50 per cent recovery time and this effect was completely reversed on rewarming

From these results it is clear that cooling has a pronounced effect on the action of depolarizing drugs The fact that a block produced by decamethonium is more or at least as much, prolonged as one produced by suxamethonium suggests that the prolongation of the effect cannot be due to inactivation of enzymes at the lower temperatures for if this were so the suxamethonium block should be the most affected This result together with the fact that substances interrupting the neuromuscular transmission by competition with acetylcholine are less affected by cooling suggests that cooling influences the

process by which a long lasting depolarization of the motor end plate interrupts neuromuscular transmission This suggestion is supported by the findings of Csapo and Wilkie (1956) who studied the effects of potassium on frog skeletal muscle and found that at low temperatures the recovery of the depolarization produced by potassium is very slow, and that a brief period of warming leads to sudden and dramatic recovery

Suxamethonium is known as a short lasting neuromuscular blocking drug From time to time, however, anaesthetists come across cases in which suxamethonium produces a prolonged apnoea Much work has been done, and many theories formed, in attempts to explain this phenomenon, but so far no satisfactory reason for it has been found If this prolonged apnoea is really due to a prolonged neuromuscular blockade of the respiratory muscles, then the initial cause of it may well be a fall in their temperature From our experiments it is obvious that even a small fall in muscle temperature can prolong the duration of a blockade produced by depolarizing drugs All these factors put together suggest to us that such a possibility should be borne in mind by the anaesthetists

REFERENCES

BIGLAND B GOETZEE B MACLAGAN J and ZAIMIS E (1958) *J Physiol* In press
BOVET D and BOVET NITTI F (1955) *Sc med ital* 3 484
BURNS B D and PATON W D M (1951) *J Physiol* **115** 41
CSAPO A and WILKIE D R (1956) *J Physiol* **134** 497
GOODMAN L S and GILMAN A (1955) *The Pharmacological Basis of Therapeutics* The Macmillan Company New York
HOLMES P E B JENDEN D J and TAYLOR D B (1951) *J Pharm exp Therap* **103** 38?
JEWELL P A and ZAIMIS E (1954) *J Physiol* **124** 417
MARSH D F (1952) *J Pharm exp Therap* **105** 299
PATON W D M and ZAIMIS E (1951) *J Physiol* **112** 311
PATON W D M and ZAIMIS E (1952) *Pharm Rev* **4** 219
ROBSON J M and KEELE C A (1956) *Recent Advances in Pharmacology* J & A Churchill Ltd London
THESLEFF S (1955a) *Acta physiol scand* **34** 218
THESLEFF S (1955b) *Acta physiol scand* **34** 386
ZAIMIS E (1951) *J Physiol* **112** 176
ZAIMIS E (1953) *J Physiol* **122** 238
ZAIMIS, E (1954) *Pharm Rev* **6** 53

XIII

Phospholipids

CYRIL LONG

THE phospholipids may be defined as those compounds which possess the general physical properties characteristic of fatty substances but which are distinguished chemically by the fact that they contain phosphorus They are substances of relatively low melting point and are soluble in several organic solvents but insoluble in water, though some of them form fairly stable emulsions with water Although the history of the phospholipids may be traced back to the closing years of the eighteenth century for a long time they did not excite any very general interest Indeed, the preliminary studies in this field are largely due to the energy of a few enthusiasts such as Thudichum the Macleans and Levene During the last twenty years however, the phospholipids have become much more popular as subjects for research and it is probable that quite soon they will be regarded as being of great significance in medicine for they are universal constituents of living cells At the present time it is not possible to do more than predict this important future for them, because there is still a great deal of fundamental work to be carried out, before the medical applications can be properly examined In what follows an attempt will be made to assess critically some of the major advances which have been made in the study of the phospholipids It will be necessary to deal with their chemical structures, the problem of their separation from one another their degradation by enzymes and their synthesis by living cells Finally it might be profitable to speculate on the possible function of one phospholipid namely lecithin in the liver

Most phospholipid biochemists would probably agree that, despite recent intensive activity, progress in the study of the phospholipids has not been as rapid as one could have wished The fact that lipids, as a class, are only soluble in certain organic solvents, is not so great a disadvantage as might be supposed, in fact, this property can be very useful in that, provided the conditions are properly chosen, lipids can be separated completely from contaminating proteins and largely from carbohydrates and other water soluble substances There are, however, two main difficulties in obtaining pure lipids First, the various lipid types have rather similar solubility characteristics, due largely to the dominating influences of the component long chain fatty acid radicals, this is so in spite of the fact that there may be clear differences in the chemical structures of the different types The problem becomes even more complex when it is realized that the presence of one type of lipid modifies the solubility properties of others For example, purified phospholipids are sparingly soluble in acetone, but in the presence of neutral fats they become much more soluble The second difficulty arises from the fact that individual phospholipid types do not occur in biological materials as single chemical substances Since their component fatty acid radicals may contain different numbers of carbon atoms or possess different degrees of saturation, their solubility properties extend over a considerable range Lecithin for example, which is the most common phospholipid, contains two fatty acid radicals in its molecule Usually one is saturated and the other unsaturated and this type of lecithin is soluble in ether However, lecithins have been isolated from animal tissues in which both fatty acid radicals are saturated and these lecithins are insoluble in ether

As a result of these two factors, there is normally a considerable overlapping of solubility properties between different lipid types In spite of these difficulties, however, until a few years ago solvent fractionation techniques were practically the only methods available for the separation of lipids and it is not surprising that the preparation of individual lipid types was extremely laborious invariably resulted in great losses of materials and normally gave products of purity not greater than

90–95 per cent and sometimes much less In an attempt to improve these separations counter-current distribution methods have been examined but there seem to be certain technical difficulties which will prevent their general application The most impressive advances have been made with the advent of chromatographic procedures and these will be dealt with later

Chemical Structures

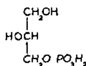
Most of the phospholipids are based on L α glycerophosphoric acid (I) [1] the two free hydroxyl groups usually being esterified with long chain fatty acids such compounds are termed phosphatidic acids (II) The phospholipid which has been most in tensively studied is lecithin (III) The phosphate group of the phosphatidic acid is esterified with choline (IV) This structure for lecithin has been firmly established both by degradative studies and by chemical synthesis (Baer and Kates, 1950, Long and Maguire, 1953 1954) Lecithin accounts for about 1–2 per cent of most animal tissues

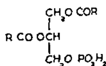
Phosphatidylethanolamine (V), previously known as 'kepha lin resembles lecithin very closely in structure but contains ethanolamine (VI) in place of choline Phosphatidylethanolamine is also present in most animal tissues to the extent of about 1–2 per cent Phosphatidylserine (VII) (Folch 1948) is based on the same plan the nitrogenous component being the amino acid serine (VIII) It is usually present in animal tissues at a lower concentration than is lecithin or phosphatidyl ethanolamine

Lysolecithin (IX) is often regarded as a breakdown product of lecithin by removal of one fatty acid radical However it occurs in appreciable quantity in egg yolk under conditions where enzymic degradation can be ruled out (Lea Rhodes and Stoll 1955)

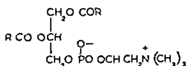
The plasmalogens (acetal phospholipids) are also based on glycerophosphate There is a great deal of uncertainty concern ing their structures due largely to the fact that they have never been isolated in pure form The structure shown (X, after Klenk and Debuch 1954 1955) is only one of several possibilities

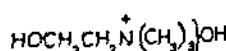
[1] Illustrations I–XVI are shown in Figure 1



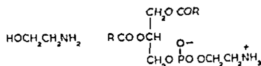


I II

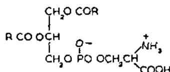


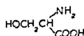


III IV

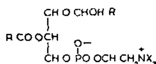


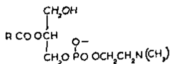
VI V



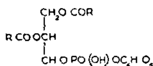


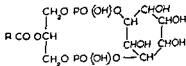
VII VIII





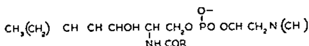
X IX



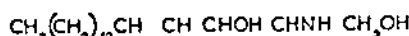


XI XII

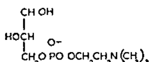
Fig 1 Structural formulae of phospholipids



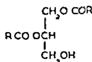
XIV



XIII



XV



XVI

FIG 1 *cont*

Ethanolamine plasmalogen is a major constituent of brain while choline plasmalogen is present in high concentration in cardiac muscle

There are at least two phosphoglycerides containing inositol Phosphatidylinositol, with the possible structure (XI) has been obtained from liver by Hawthorne (1955), while a sub stance with one more phosphate group and one less fatty acid radical, diphosphoinositide (Folch 1949) of possible structure (XII), is a constituent of brain

A well known phospholipid containing no glycerol but based instead on sphingosine (XIII) is sphingomyelin (XIV) It is present in large amount in brain

Perhaps the best comment on this catalogue of the various phospholipid types is that with few exceptions there is a lesser or greater degree of uncertainty concerning their chemical structures This is largely due to the fact that to date satis factory methods are not available for the isolation of most of the various types in pure form This applies especially to the plasmalogens and the inositol phospholipids

SEPARATION OF LIPIDS FROM ONE ANOTHER

At the present time there is a surge of activity in several centres in an attempt to effect complete and quantitative separations

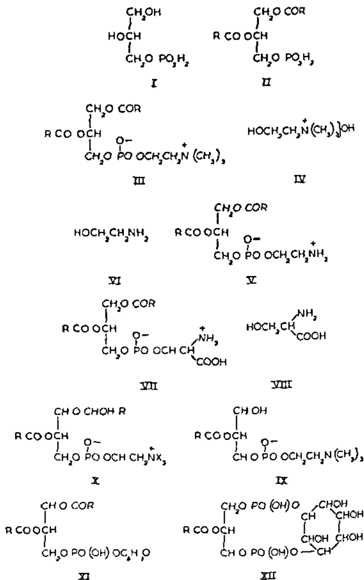


Fig 1 Structural formulae of phospholipids

Enzymic degradation As far as enzymic breakdown is concerned it will be noted that the lecithin molecule contains four ester linkages and, as would be expected, these four linkages are hydrolysed by four separate enzymes Figure 3 shows their sites of action These enzymes have been studied in detail and

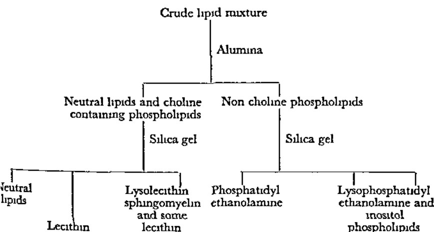


FIG 2 The chromatographic separation of egg yolk phospholipids Method of Rhodes and Lea (1957)

are known as phospholipases *A B C* and *D* The term lecithinases which was originally used for these enzymes has largely been abandoned owing to their known lack of specificity Experiments carried out several years ago (Fairbairn 1945 Kahane and Levy, 1945) suggested that the intracellular enzymes of animal tissues brought about a complete hydrolysis of endogenous lecithin when the tissues were homogenized and incubated, so that there subsequently arose a general feeling that animal tissues with their mixed battery of enzymes were unpromising materials for the study of individual enzyme activities During the last ten years therefore, much effort has been devoted to the study of single enzymes which are found in biological materials other than animal tissues In the following paragraphs the action of phospholipases on lecithin only will be considered

of the various types of lipids The lipids of egg yolk have received special attention, because the mixture is less complex than those obtained from animal tissues, being virtually free from plasmalogens and phosphatidylserine One of the most useful observations is that of Hanahan, Turner and Jayko (1951), who have shown that chromatography of mixed lipids on an alumina column will quantitatively retain non choline phospholipids from a chloroform methanol solution and permit the neutral lipids and choline containing phospholipids to pass through with the solvent front Rhodes and Lea (1957) have extended this method and have been able to elute the non choline phospholipids from the column with a more polar solvent (ethanol chloroform water) Thus, in one stage, a quantitative separation of egg yolk lipids into two distinct fractions with no overlapping of components has been achieved The two fractions may then be further separated by chromatography on a silica gel column (Figure 2) The first fraction from the alumina column is separated into neutral lipids, pure lecithin, and a mixture of lysolecithin and sphingomyelin contaminated with some lecithin The non choline phospholipid fraction is separated into pure phosphatidylethanolamine and a mixture of lysophosphatidyl ethanolamine and inositol phospholipids If the more complex phospholipid mixtures from animals tissues were to be subjected to this procedure however, it seems likely that the lecithin would be contaminated with choline plasmalogen, while the phosphatidylethanolamine would be contaminated with both ethanolamine plasmalogen and phosphatidylserine Nevertheless, these separations are a great advance on earlier methods They have the virtue of being quantitative and applicable to small quantities and it may be hoped that still further separations will result from the use of other types of column

Phospholipid Biochemistry

Phospholipid biochemistry may be considered under two main headings, namely breakdown and synthesis by enzymes Lecithin has been most widely studied in this connection largely because it is the most readily available phospholipid and because its structure is firmly established

tion by animal tissues may be reconsidered In the first place, it seems likely that phospholipases *C* and *D* are not implicated the former being a bacterial enzyme and the latter a plant enzyme On the other hand there is evidence that lecithin is normally degraded in animal tissues by the successive actions of phospholipases *A* and *B* While it is true that lysolecithin the product of phospholipase *A* activity, does not appear to accumulate in animal tissues, an active phospholipase *A* has in fact been extracted from pancreas (Noguchi, 1944) Phospholipase *B* is present in liver tissue, as shown by Dawson (1956b) Furthermore, glycerylphosphorylcholine was found to accumulate in autolysing pancreas and intestine (Schmidt, Hershman and Thannhauser, 1945) and tracer experiments with liver tissue by Dawson (1955) have shown that it must be a degradation product of lecithin rather than a precursor It is probable that the phospholipase *B* activity of tissues always exceeds their phospholipase *A* activity, so that toxic amounts of lysolecithin do not accumulate Since liver also contains an enzyme which hydrolyses glycerylphosphorylcholine into glycerophosphate and choline (Dawson 1956a), it seems clear that the enzymic breakdown of lecithin in liver and possibly in other animal tissues can be represented by the following pathway lecithin $\rightarrow$ lysolecithin $\rightarrow$ glycerylphosphorylcholine $\rightarrow$ glycerophosphate + choline

Enzymic synthesis In studies of the biosynthesis of lecithin, the first clear indications were derived from experiments on animals to which radioactive inorganic phosphate had been administered The lecithin and water soluble phosphate esters related to lecithin were separated from the tissues and the specific radioactivity/time curves were then examined by criteria which distinguish precursors from non precursors (Zilversmit Entenman and Fishler, 1943) In this way it has been shown by Popjak and Muir (1950) that α glycerophosphate is a precursor of lecithin in the liver, however, other observations on liver by Kornberg and Pricer (1952) have shown that phosphorylcholine is also a precursor of lecithin This result, that both glycerophosphate and phosphorylcholine can be precursors of lecithin in liver is rather surprising especially when it

The venom of several species of snake, including the cobra, rattlesnake and especially the black cottonmouth moccasin, contains a very potent phospholipase *A* This enzyme has been shown to split off the fatty acid radical from the α position of the lecithin molecule, the product being lysolecithin (IX,

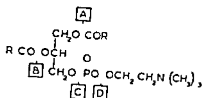


Fig 3 The enzymic hydrolysis of lecithin The letters denote the site of attack of phospholipases *A B C* and *D*

Hanahan, 1954, Long and Penny, 1954) Lysolecithin is a haemolytic substance and is responsible in part for the ill effects following snake bite Lysolecithin can serve as substrate for another enzyme phospholipase *B*, which is present in large amount in the mycelial felts of *Penicillium notatum* (Fairbairn, 1948) As a result the remaining fatty acid radical on the β position is split off, giving glycerylphosphorylcholine (XV) This substance has no fatty acid radicals and is therefore water soluble

Enzymes are also known which attack the other end of the lecithin molecule For example, phospholipase *C*, which is present in culture filtrates of the strict anaerobe *Clostridium welchii* and related species (Macfarlane and Knight, 1941, Macfarlane 1948) splits off phosphorylcholine leaving an α β diglyceride (XVI Long and Maguire, 1954) This enzyme has been identified with the α toxin of *Clostridium welchii* and is the factor causing the haemolysis associated with gas gangrene, the enzyme presumably attacking the lecithin present in the erythrocyte membrane

Phospholipase *D* has been found in several plants and seems to be present in highest concentration in Savoy cabbage and Brussels sprouts (Davidson Long and Penny, 1956) It removes choline from lecithin giving the corresponding phosphatidic acid (II)

Having obtained this detailed knowledge of the activities of the individual phospholipases the problem of lecithin degrada

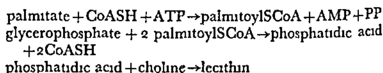
have been considered, but ultimately, every phospholipid bio chemist must ask What is the nature of the primary function of phospholipids in living cells? At the present time there is no answer to this question which is acceptable to all workers, although most would agree that phospholipids are probably concerned in some way with the permeability of cells and of the cytoplasmic particles of cells and that their metabolism is associated with fatty acid metabolism in a way that is not understood

Since several types of phospholipid exist, it is reasonable to infer that there are at least an equal number of functions, more over, it cannot be assumed that any individual type of phos pholipid will have the same function in different kinds of cells although the differences may well be quantitative rather than qualitative If the function of even one single type of phospholipid in one kind of cell were understood a sound beginning to the elucidation of this difficult problem would have been made The following section presents some speculations on a possible role for lecithin in the metabolism of the liver

Although there is a very considerable literature on the role of lecithin in the liver the bulk of it is difficult to interpret A great deal of work has been done with intact animals, a fact which in itself introduces uncertainty because it is well established that both triglycerides and lecithin are continually passing from the plasma to the liver while lecithin also passes from the liver to the plasma These transfers are difficult to allow for Then again, many workers owing to the inadequacy of available methods of separation have expressed their results in terms of total phospholipid Since this is a heterogeneous fraction, the observations are susceptible to more than one interpretation Even when experiments have been carried out with homogenized liver preparations and when the lecithin has been isolated in pure form by chromatography, there can still be some uncertainty owing to the fact that the lecithin will be derived from different sub cellular fractions all of which may not behave in an identical manner In short it seems that the only reliable results are those which have been obtained from cytologically homogeneous cell fractions from which the lecithin has been obtained in chemically pure form

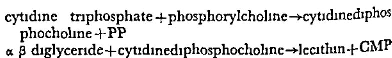
is borne in mind that glycerylphosphorylcholine is not (Dawson, 1995) The only logical conclusion is that there must be two separate pathways for the biosynthesis of lecithin in liver, and these two mechanisms have now been fully established by the work of Kennedy in Chicago and of Kornberg in Bethesda during the last few years

Kennedy (1953) showed that rat liver mitochondria converted α glycerophosphate into what appeared to be a phosphatidic acid Almost simultaneously and independently, Kornberg and Pricer (1953) showed that guinea pig liver microsomes catalysed the esterification of one molecule of glycerophosphate by two molecules of palmitic acid, that the product was in fact a phosphatidic acid and that the coenzyme A derivative of palmitic acid was an intermediate Kennedy (1954) showed that choline could be incorporated into lecithin by rat liver mitochondria, that ATP was required, and that neither phosphorylcholine nor glycerylphosphorylcholine were intermediates The complete process can be represented by the following scheme



palmitate + CoASH + ATP → palmitoylSCoA + AMP + PP
glycerophosphate + 2 palmitoylSCoA → phosphatidic acid + 2CoASH
phosphatidic acid + choline → lecithin

The phosphorylcholine pathway was clarified by Kennedy and Weiss (1955), who showed that cytidinediphosphocholine was an intermediate The reaction pathway is probably



cytidine triphosphate + phosphorylcholine → cytidinediphosphocholine + PP
α β diglyceride + cytidinediphosphocholine → lecithin + CMP

Although the acceptor molecule has not been definitely identified as an α β diglyceride, Rodbell and Hanahan (1955) showed that αβ dipalmitin stimulated the rate of incorporation of phosphorylcholine into lecithin by guinea pig liver mitochondria

A possible Function for Lecithin in Liver Metabolism

So far, the chemistry and enzymology of the phospholipids

mately choline is set free which can in turn be converted in to phosphorylcholine which can then react with more di glyceride The cycle is then repeated and with each turn, *one molecule of diglyceride yields its fatty acids for oxidation*

Some supporting evidence in favour of this scheme can be

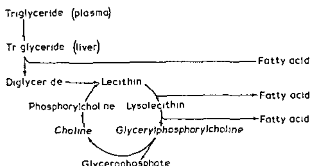


FIG 4 The phosphorylcholine cycle A possible role for lecithin in the metabolism of triglycerides in the liver

found in a paper by Rodbell and Hanahan (1955a) These workers found that the oxidation of the fatty acid radicals of the endogenous lipids of guinea pig liver mitochondria was stimu lated by the addition of lysolecithin glycerylphosphoryl choline and especially by phosphorylcholine, when used in catalytic amounts

The choline formed in this cycle will partly be destroyed by the action of the liver choline oxidase in those species which possess this enzyme Thus the rat whose liver contains choline oxidase is well known to require a constant supply of choline or choline precursors in the diet In their absence the cycle would cease to function and triglycerides would accumulate in the liver This could explain the production of fatty livers in rats fed a diet deficient in choline On the other hand the liver of the guinea pig contains no choline oxidase and it has been observed repeatedly that guinea pigs do not suffer from fatty livers when fed a diet deficient in choline

With regard to the other pathway, *the glycerophosphate* almost certainly arises by the reduction of dihydroxyacetone phosphate an intermediate in the glycolysis of carbohydrate

There are only three well established facts about lecithin metabolism in the liver, namely the pathway of degradation and the two pathways of biosynthesis, and these are the only facts which may legitimately be used in an attempt to establish a role for lecithin in liver function. So far, the two biosynthetic pathways have been distinguished in terms of the phosphate group. However, it is probably more useful in the present discussion to distinguish them in terms of the glycerol portion of the molecule. In the first pathway, the glycerol is already esterified with fatty acid radicals before the attachment of the phosphate, whereas in the second the glycerol is esterified to phosphate before the addition of the fatty acid radicals. In stating the problem in this way, we are led to a search for the origin of the diglyceride and of the glycerophosphate.

There is at present no very good evidence for the occurrence of free glycerol in liver, so that the diglyceride is unlikely to be formed by the partial esterification of glycerol with fatty acids. On the other hand, there is always an appreciable amount of triglyceride in liver, probably originating from the fat depots and transported to the liver via the plasma, and a partial hydrolysis of triglyceride could produce the necessary diglyceride. This idea is supported by the finding of an enzyme in liver which hydrolyses esters of long chain fatty acids (Copenhaver, Stafford and McShan, 1950), and if this enzyme has an action pattern similar to that of the well known pancreatic lipase (Mattson and Beck, 1955), the diglyceride formed will have the $\alpha\beta$ configuration, which is the isomer required for reaction with phosphorylcholine. In the normal animal the fatty acid radicals of liver triglyceride are readily oxidized, but before this can happen the three fatty acids must be released from the triglyceride. The lipase enzyme is unable to do this by itself, again assuming that the liver and pancreatic enzymes are similar, because lipase cannot remove the fatty acid radical from the β position of the triglyceride, and in fact readily liberates only one fatty acid radical. The scheme shown in Figure 4 indicates a possible way in which these three fatty acid radicals may be made available for oxidation. Liberation of two of these fatty acids is associated with the breakdown of the lecithin, and ulti

these ideas are untenable, however, then they will have to be discarded, a fate suffered by many earlier theories If and when this happens, however we shall surely be approaching nearer to the truth, and this after all is the ultimate aim of scientific research

REFERENCES

BAER E and KATES M (1950) *J Amer chem Soc* **72** 942

COPENHAVER J H STAFFORD R O and McSHAN W H (1950) *Arch Biochem* **26** 260

DAVIDSON F M LONG C and PENNY I F (1956) In *Biochemical Problems of Lipids* Ed Popjak and Le Breton Butterworth s Scientific Publications London

DAWSON R M C (1955) *Biochem J* **59** 5

DAWSON R M C (1956a) *Biochem J* **62** 689

DAWSON R M C (1956b) *Biochem J* **64** 192

FAIRBAIRN D (1945) *J biol Chem* **157** 654

FAIRBAIRN D (1948) *J biol Chem* **173** 705

FOLCH J (1948) *J biol Chem* **174** 439

FOLCH J (1949) *J biol Chem* **177** 505

HANAHAN D J (1954) *J biol Chem* **207** 879

HANAHAN D J TURNER M B and JAYKO M E (1951) *J biol Chem* **192** 623

HAWTHORNE J N (1955) *Biochem J* **59** 11

KAHANE E and LEVY J (1945) *Bull Soc chim Biol* **27** 544 588

KENNEDY E P (1953) *J biol Chem* **201** 399

KENNEDY E P (1954) *J biol Chem* **209** 525

KENNEDY E P and WEISS S B (1955) *J Amer chem Soc* **77** 250

KLENK E and DEBUCH H (1954) *Hoppe Seyl Z* **296** 179

KLENK E and DEBUCH H (1955) *Hoppe-Seyl Z* **299** 66

KORNBERG A and PRICER W E (1952) *Fed Proc* **11** 24?

KORNBERG A and PRICER W E (1953) *J biol Chem* **204** 345

LEA C H RHODES D N and STOLL R D (1955) *Biochem J* **60** 353

LONG C and MAGUIRE M F (1953) *Biochem J* **54** 612

LONG C and MAGUIRE M D (1954) *Biochem J* **57** 223

LONG C and PENNY I F (1954) *Biochem J* **58** xv

MACFARLANE M G (1948) *Biochem J* **42** 587 590

MACFARLANE M G and KNIGHT B C G J (1941) *Biochem J* **35** 884

MATTSON F H and BECK L W (1955) *J biol Chem* **214** 115

NOGUCHI S (1944) *J Biochem Japan* **36** 113

POPJAK G and MUIR H (1950) *Biochem J* **46** 103

RHODES D N and LEA C. H (1957) *Biochem J* **65** 526

Fatty acid biosynthesis in liver takes place by the condensation of several molecules of acetylcoenzyme A to give long chain fatty acylcoenzyme A, presumably by a pathway resembling the fatty acid oxidation spiral in reverse The acetylcoenzyme A molecules are also derived from carbohydrate Since there is a limited

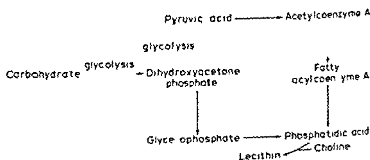


FIG 5 A possible function for the glycerophosphate pathway of lecithin synthesis in the conversion of carbohydrate into fat in the liver

quantity of coenzyme A in liver cells, the biosynthesis of long chain fatty acylcoenzyme A would soon cease owing to the immobilization of coenzyme A However if the fatty acyl coenzyme A molecules can condense with an acceptor molecule such as glycerophosphate the coenzyme A will be regenerated The phosphatidic acid thus formed could then react with choline to give lecithin If this scheme, which is shown in Figure 5 is correct then it follows that in the early stages of long chain fatty acid synthesis in liver using tracer acetate, the fatty acid radicals of the lecithin will be labelled more strongly than those of triglyceride and this has been shown to be the case by several workers (see e g Tove Andrews and Lucas 1956) The fate of the lecithin so formed is at present uncertain There is no known mechanism for the conversion of lecithin into triglycerides although it is very likely that such a conversion does take place either in the liver itself or in the extrahepatic tissues to which the lecithin is transported via the plasma

These two schemes for lecithin function in the liver which have been tentatively suggested are admittedly very speculative Many of the observations in the literature can be understood if these pathways are correct If further evidence shows that

XIV

The Synthesis and Degradation of Polysaccharides

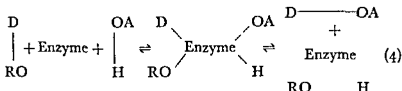
W J WHELAN

If we define the term polysaccharide to mean the naturally occurring sugar polymers containing ten or more mono saccharide residues per molecule then the known examples of their enzymic synthesis by cell free enzyme systems are very few In the plant and animal kingdoms the only known systems are those responsible for the synthesis of starch and glycogen respectively, and even with the inclusion of the microbial enzymes the field is restricted to the synthesis of three types of linkage the α 1 4 and α 1 6 glucopyranosidic linkages and the β 2 6 fructofuranosidic linkage All these systems were dis covered before 1950 [1]

Despite this lack of progress in the discovery of new systems the last decade has been a period of great advances in our under standing of the mechanism of the enzymic synthesis of poly saccharides This is because of the rapid progress in the related field of oligosaccharide synthesis Oligosaccharides may be considered to comprise the range of sugar polymers containing

[1] Since this review was prepared three new typ s of polysaccharide synthesis have been reported Each one utilizes a uridine diphosphate sugar compound as the donor substr te ( ) Hyaluronic acid synthesized by an enzyme system from Rous sarcoma of ch ckens using uridine diphosphate *N* acetylglucosamine (Glaser L (1957) *Diss Abstr* **17** 29) (i) Ch t n is synthesized by a particulate enzyme f om *Neuro pora crass* with uridine diphosphate *N* acetylglucosam ne (Glaser L and Brown D H (1957) *Biochem biophys Acta* **23** 449) (iii) Cellulose is synthe sized by a cell free part culate system from *A tobacter xylinum* using uridine diphos phate glucose as donor substrate (Glaser L (1957) *Biochem biophys Acta* **25** 436 see also Colvin J R (1957) *A ch Biochem Biophys* **70** ?94 and Greathouse G A (1957) *J Amer chem S* **79** 4503)

RODBELL M and HANAHAN D J (1955a) *J biol Chem* **214** 595
RODBELL M and HANAHAN D J (1955b) *J biol Chem* **214** 607
SCHMIDT, G HERSHMAN B and THANNHAUSER (1945) *J biol Chem* **161** 523
TOVE S B ANDREWS J S and LUCAS H L (1956) *J biol Chem* **218** 275
ZILVERSMIT, D B ENTENMAN, C and FISHLER M C (1943) *J gen Physiol* **26** 325



An exception to this generalization is the breakdown of hyaluronic acid by certain bacterial enzymes, with the formation of an unsaturated disaccharide (Linker, Meyer and Hoffman, 1956) This reaction does not seem to involve an acceptor substrate

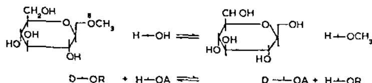


FIG 1 Hydrolysis of ^{18}O labelled methyl β glucoside by almond emulsin Glucosyl-oxygen scission also occurs in hydrolysis by hydrochloric acid and when methyl α-glucoside is hydrolysed by yeast α glucosidase or acid (Bunton *et al* 1954)

Examples which illustrate the generalization are the hydrolysis of methyl β glucoside by almond β glucosidase (emulsin) (Figure 1) and the synthesis of α 1 4 glucosidic linkages by muscle phosphorylase (Figure 2) In these and other cases it is known that the enzyme splits the glycosyl oxygen bond in the donor substrate and transfers the glycosyl radical not the glycosidic (glycosyloxygen) radical (Cohn 1949 Bunton Lewis Llewellyn Tristam and Vernon 1954, Koshland and Sein 1954) The enzymic transfer of sugar residues is therefore known as transglycosylation

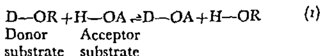
SYNTHESIS OF STARCH IN PLANTS

It would be more appropriate in this series of lectures to deal first with the synthesis of glycogen in animals but there are fewer unsolved problems pertaining to the similar enzyme system responsible for the synthesis of starch and this is therefore easier to describe

from two to nine monosaccharide residues per molecule[1] Whereas ten years ago we knew more about polysaccharide than oligosaccharide synthesis, today the position is reversed but the two fields are so intimately related that our knowledge of polysaccharide synthesis has inevitably increased The main factor in these advances has been the chromatographic method in all its forms, but particularly paper chromatography, and to a lesser extent ionophoretic and isotopic tracer techniques If there were some method, chromatographic or otherwise, of detecting polysaccharide synthesis as readily as oligosaccharide synthesis can now be followed, then the former field would undergo the same rapid development It seems likely that such methods will be forthcoming

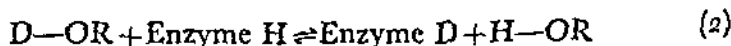
The Mechanism of Carbohydrase Action

Almost all synthetic and degradative reactions of polysaccharides as catalysed by enzymes can be expressed in terms of a reaction between a donor and an acceptor substrate as shown in the following equation

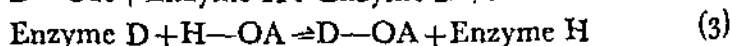


$$\underset{\text{Donor substrate}}{\text{D—OR}} + \underset{\text{Acceptor substrate}}{\text{H—OA}} \rightleftharpoons \text{D—OA} + \text{H—OR} \qquad (1)$$

In this equation the group D must be a carbohydrate residue, R and A may also be carbohydrate in nature but need not necessarily be so It has been suggested (see Barker and Bourne, 1953) that equation (1) can be split into two equations which show how the enzyme participates in the reaction



$$\text{D—OR} + \text{Enzyme H} \rightleftharpoons \text{Enzyme D} + \text{H—OR} \qquad (2)$$



$$\text{Enzyme D} + \text{H—OA} \rightleftharpoons \text{D—OA} + \text{Enzyme H} \qquad (3)$$

An alternative suggestion (Jermyn 1957) is as follows

[1] The distinction between oligo- and poly saccharides (Whistler and Smart 1953) is not intended to be precise but the upper limit of oligosaccharide size is also near the limit of the ability of paper chromatography to separate individual sugars and the defined range of oligosaccharide size includes most of the products formed by the so-called oligosaccharide synthesizing systems

Historically speaking the first clue to the mechanism of starch synthesis came from the demonstration in 1937 by Cori and others that rabbit muscle and liver contained an enzyme system capable of effecting a reversible conversion of α D glucopyranose 1 phosphate into glycogen This was followed by Hanes s finding of a similar enzyme system in peas and potatoes (Hanes 1940a, b) and thereafter the developments in both fields have kept in step Hanes made two important contributions He recognized that synthesis of starch did not take place unless preformed chain molecules of α 1 4 linked glucose units ('primer) were added to the sugar phosphate and enzyme It was presumed, and all later evidence confirms this, that synthesis takes place by the apposition of α glucose units derived from the ester (donor substrate), to the non reducing chain ends of the acceptor substrate the primer molecules (Figure 2) The position of attachment is at the C 4 hydroxyl of the non reducing end glucose unit The reaction is freely reversible and in the presence of excess primer the equilibrium is influenced only by the concentrations of inorganic and ester phosphates (see Whelan (1955) for details and for Baum and Gilbert s method of crystallizing potato phosphorylase) The smallest molecule which will act as primer is maltotriose but this is very inefficient by comparison with maltotetraose and larger maltodextrins (Weibull and Tiselius 1945 French and Wild 1953 Whelan and Bailey, 1954) Some dextrans (α 1 6 linked polyglucoses) have been reported to prime the reaction but only α glucose 1 phosphate of the many esters tested is effective as a donor substrate (see Whelan, 1955) The second contribution made by Hanes was the demonstration that the purified enzyme (phosphorylase) did not synthesize starch but a polysaccharide having the properties of one of the two components of starch namely amylose In particular, the synthetic polysaccharide had the intense blue iodine stain and complete degree of β amylolysis to maltose characteristic of the natural material, the molecule of which is an essentially unbranched chain of several thousand α 1 4 linked glucose units The other starch component amylopectin has a weak purple iodine stain and only about half of the molecule can be attacked by β amylase (see later) The enzymic

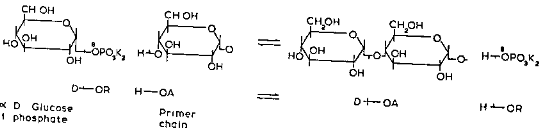


FIG 2 The mechanism of the unit step in the synthesis of amylose by rabbit muscle phosphorylase as shown by the use of ^{18}O labelled glucose 1 phosphate Sucrose phosphorylase (Doudoroff Barker and Hassid 1947) and hydrogen ions also cause glucosyl oxygen scission of the Cori ester but acid and alkaline phosphatases hydrolyse it by glucosidic phosphorus scission (Cohn 1949)

units per molecule, but this may be due to the presence of amylolytic impurities in the Q enzyme preparation (Nussenbaum and Hassid 1952) Q Enzyme does not seem to be able to reverse the conversion of amylose into amylopectin but it may perhaps be able to reposition the α 1 6-links It is known

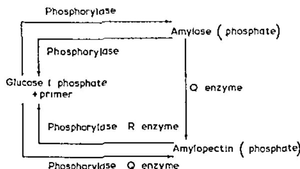


Fig 4 The enzymic synthesis and degradation of amylose and amylopectin

that the donor substrate must be at least 40 units long before Q enzyme can exert a rapid action but the requirements of the enzyme towards the acceptor substrate have not been defined (Peat Whelan and Bailey 1953) The protozoal Q enzyme of *Polytomella coeca* will utilize maltose as an acceptor (Barker Bebbington and Bourne, 1953) Amylopectin is also synthesized when Q enzyme is added to a mixture of phosphorylase primer and glucose 1 phosphate If only small amounts of branching enzyme are added and the reaction is not allowed to go to completion, products intermediate in structure between amylose and amylopectin can be isolated (Barker Bourne Peat and Wilkinson 1950) The interrelationships between phosphorylase and Q enzyme are shown in Figure 4

There is a third enzyme in the potato which probably functions alongside phosphorylase and Q enzyme in the formation of starch molecules This is D enzyme (Peat, Whelan and Rees 1956) which transfers chain segments of α 1 4 linked glucose units from one starch molecule (donor) to another such molecule or to glucose or a variety of related acceptor sugars (Peat Whelan and Jones 1957 Walker and Whelan 1957) The

synthesis of this latter polysaccharide was achieved by Haworth Peat and Bourne (1944) by means of Q enzyme from the potato This converts amylose into amylopectin (Barker, Bourne and Wilkinson, 1950) and has been crystallized by Baum and Gilbert (1953) The net effect of Q enzyme action is that about

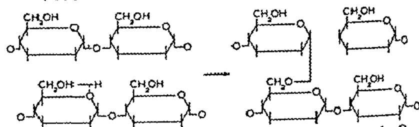


Fig 3 Unit step in the synthesis of amylopectin from amylose by Q enzyme (It is not yet proved that glycosyl-oxygen scission takes place)

5 per cent of the α 1 4 linkages of amylose are converted into α 1 6 linkages The enzyme is presumed to act by severing a portion of an amylose chain (donor substrate) and attaching the severed portion through its reducing end to the primary hydroxyl group of another chain (acceptor substrate) (Figure 3) The branching action seems to cease when the proportion of non reducing chain ends has increased from one in several thousand to one in twenty as in natural amylopectin The resultant amylopectin molecule is a tree like structure containing the unit chains (average DP about 20) cross linked through the α 1 6 bonds in random fashion as envisaged by Meyer and Bernfeld (1940) and proved for natural amylopectin by Peat Whelan and Thomas (1952 1956) The branch linkages are so situated that β amylase which attacks starch type molecules from the non reducing chain ends splitting off maltose in its progress along the chain is able to remove only about half of the molecule as the disaccharide before its endwise action is stopped by the approach of the enzyme to the points of branching, which it cannot circumvent Calculations of the lengths of the outer and inner chains of amylopectin and glycogen have been made by Manners (1953 1955) (see Figure 7) The synthetic material has a lower degree of polymerization than natural amylopectin, which contains several thousand glucose

polysaccharide floridean starch (from *Dilsea edulis*) also contains the α 1 3 linkage (Peat Turvey and Evans, 1957)

There is a great deal of circumstantial evidence of the interdependence of the metabolism of sucrose, the first free sugar to be formed in photosynthesis (Calvin, 1956) and glucose

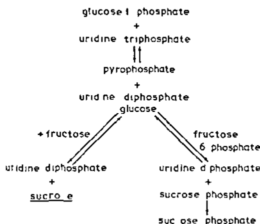


FIG 6 The enzymic synthesis of sucrose from glucose 1 phosphate

1 phosphate, the starch precursor The enzymic interrelation ship of these two substances has recently been established (Buchanan 1953 Cardini Leloir and Chiriboga 1955 Leloir and Cardini 1955, Bean and Hassid 1955 Burma and Morti mer 1956) It appears that there are two routes connecting glucose 1 phosphate and sucrose (Figure 6) The key inter mediate is uridine diphosphate glucose (Figure 5) This is formed in the reaction between uridine triphosphate and glu cose 1 phosphate Uridine diphosphate glucose can then react either with fructose to yield sucrose and uridine diphosphate or with fructose 6 phosphate to form sucrose phosphate and uri dine diphosphate The sucrose phosphate then yields sucrose by the action of a phosphatase The enzyme system has been detected in a wide variety of plant parts

The phosphorolysis of starch The degradation of starch may be accomplished by the hydrolytic amylases (see later) or by

linkage synthesized is the same as that broken, i e the α 1 4 bond When acting on maltotriose (the smallest donor substrate), glucose and maltodextrins larger than maltotriose are formed, but neither the number of α 1 4 glucosidic linkages nor the average degree of polymerization of all the molecules pre

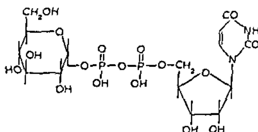


FIG 5 Uridine diphosphate glucose

sent changes during this action The role of D enzyme is perhaps to move fragments of starch chains from one site to another Whelan (1957) has discussed possible *in vivo* functions of D enzyme in regulating the supply of primer for phosphorylase, in assisting in the phosphorolysis of starch (i e conversion into glucose 1 phosphate) and in a scheme to explain the simultaneous deposition of amylose and amylopectin in the starch granule D Enzyme can presumably synthesize amylose from the small maltodextrin oligosaccharides such as maltotriose It is known that the addition of glucose to a mixture of D enzyme and maltodextrins causes the average degree of polymerization of the oligosaccharides to decrease (Peat, Whelan and Rees, 1956) Presumably the continuous *removal* of glucose would have the opposite effect and cause the oligosaccharides to increase in length, as in the analogous case of amylomaltase (see below)

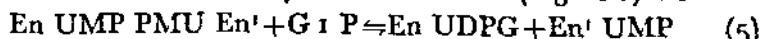
It has recently been discovered that waxy maize starch, which is virtually free from amylose contains a small proportion of α 1 3 linkages in addition to the previously recognized α 1 4 and α 1 6 linkages (Wolfrom and Thompson 1956) The enzyme responsible for the synthesis of these linkages is not known but it has been suggested that Q enzyme may be responsible, as well as for the synthesis of α 1 6 linkages (Pazur, Budovich and Tipton 1957) The amylopectin like seaweed

Synthesis and Degradation of Glycogen in Animals

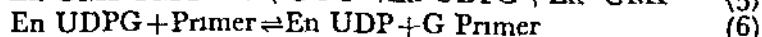
Glycogen is a polyglucose having the same α 1 4 and α 1 6 linkages and the same tree like structure as amylopectin (Larner, Illingworth, Cori and Cori 1952), but the unit chains of glycogen are only about half the length of the amylopectin chains, containing from 10–14 units although variations outside this range have been noted (Abdel Akher and Smith, 1951 Manners, 1955, Manners and Archibald, 1957) Animals do not apparently store an amylose like polysaccharide

The glycogen synthesizing and degrading enzymes correspond in type to those of plants but differ from them in many points of detail Most investigators have chosen rabbit liver and muscle as the sources of the glycogen metabolizing enzymes These organs contain a phosphorylase a branching enzyme [amylo (1 4→1 6) transglucosidase] and a debranching enzyme (amylo 1 6 glucosidase) corresponding in function to plant phosphorylase, Q enzyme and R enzyme respectively

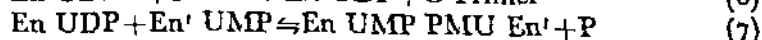
Rabbit muscle and other animal phosphorylases exist in two forms known as phosphorylases *a* and *b* Methods of crystallizing phosphorylases are described by Cori, Illingworth and Keller (1955) and Sutherland (1955) The two forms are interconvertible Phosphorylase *b* is inactive in the absence of adenylic acid but *a* loses only about 30 per cent of its activity if adenylic acid is removed Recent work on the structures of these two enzymes has been reviewed by Korkes (1956) who suggests that phosphorylase *a* consists of two protein moieties (En and En') linked to each other through two molecules of uridine mono phosphate (UMP) and that the transfer of glucosyl units from glucose 1 phosphate (G 1 P) proceeds via a uridine diphosphate glucose enzyme complex (En UDPG) in a manner somewhat analogous to the transfer of glucosyl units from glucose 1 phosphate to fructose in the synthesis of sucrose (Figure 6) i e



$$\text{En UMP PMU En}' + \text{G 1 P} \rightleftharpoons \text{En UDPG} + \text{En}' \text{ UMP} \quad (5)$$

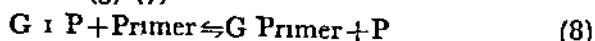


$$\text{En UDPG} + \text{Primer} \rightleftharpoons \text{En UDP} + \text{G Primer} \quad (6)$$



$$\text{En UDP} + \text{En}' \text{ UMP} \rightleftharpoons \text{En UMP PMU En}' + \text{P} \quad (7)$$

The sum of equations (5) (7) is



$$\text{G 1 P} + \text{Primer} \rightleftharpoons \text{G Primer} + \text{P} \quad (8)$$

phosphorylase In the same way that phosphorylase builds up starch chains by apposition of glucosyl residues (from glucose 1 phosphate) to the non reducing chain ends of the primer molecules, its degradative action on starch (donor) in presence of excess inorganic phosphate (acceptor) involves the removal of glucosyl units from the non reducing ends This endwise action ceases when the enzyme approaches a glucose unit involved in substitution additional to C 1 and C 4, e g the 6-substituted glucose unit at the branch points of amylopectin The conversion of amylopectin into glucose 1-phosphate is only about 40 per cent (Bourne, Sitch and Peat, 1949) The removal of the 1 6-links of amylopectin, which allows of a more complete degradation by phosphorylase or other endwise attacking enzymes such as β amylase (see below), is accomplished by R enzyme, found in potatoes and broad beans (Hobson, Whelan and Peat, 1951) This enzyme hydrolyses the amylopectin α 1 6-branch linkages but does not act on any and every type of α 1 6 glucosidic link For example, it does not hydrolyse the α 1 6 linked glucose disaccharide isomaltose, nor the poly glucose dextran, containing the same linkage It seems that the substrate for R enzyme must also contain α 1 4 glucosidic linkages (see Whelan, 1953) and steric factors influencing the enzyme action are also evident R-Enzyme does not attack the α 1 6 linkages in glycogen, which are much closer to each other than in amylopectin (see below) These links do, however, become accessible to the enzyme if the glycogen is first split into small fragments by α amylase or by acid (Peat, Whelan, Hobson and Thomas, 1954)

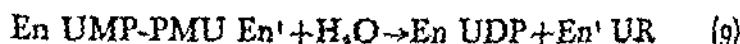
Observations on the degradation of amylose by phosphorylase are conflicting Hestrin (1949) and Peat, Thomas and Whelan (1952) found only 70 per cent conversion by muscle and potato phosphorylases respectively, and the latter workers showed that a soya bean preparation (Z enzyme), believed to be a β glucosidase rendered the amylose susceptible to more complete phosphorolysis On the other hand Baum Gilbert and Scott (1956) suggest that the obstacles to the phosphorolysis of amylose are artefacts introduced by aerobic oxidation of amylose

order to explain conflicting information about the chain length it was assumed that the discrepancy may be due to the breaking off of the long [synthetic] chains (Cori *et al* 1945) If this happened and the glycogen primer were set free then it should continue to prime the reaction whereas with small amounts of glycogen, synthesis ceases long before the phosphate equilibrium is reached (Cori *et al*, 1945) This question seems worthy of further investigation as also are the observations already mentioned that glycogen is a primer but linear chains of the same length are not while longer linear chains (amylose) are effective The suggestion is that the size of the whole molecule as well as its content of non reducing end groups is a controlling factor (cf Swanson and Cori, 1948)

The branch points in glycogen are synthesized by the branching enzyme amylo (1 4→1 6) transglucosidase from rabbit heart and liver (Cori and Cori 1943 Larner, 1955) When acting in conjunction with phosphorylase on glucose 1 phosphate and glycogen primer the enzyme gives rise to glycogen Unlike Q enzyme the branching enzyme does not act on amylose (Cori and Cori 1943) but will introduce additional branch linkages into an already branched molecule i e it will act on glycogen outer chains if these are more than 6 units long (Larner, 1953), presumably transferring the severed portions to accepting primary hydroxyl groups in the same or in other molecules Unlike *Polytomella coeca* Q enzyme (see above) it does not transfer glucose chains to maltose or other simple sugars (Larner and Uwah 1956)

Phosphorolysis of glycogen In order to effect the complete phosphorolysis of glycogen it is necessary as with amylopectin to remove the points of branching which impede the endwise glucose removing action of phosphorylase The animal counterpart of the plant debranching enzyme (R enzyme) is amylo 1 6 glucosidase, from rabbit muscle (Cori and Larner 1951, Illingworth Larner and Cori 1952 Larner *et al* 1952 Cori 1955, Larner and Schliselfeld 1956) Unlike R enzyme amylo 1 6 glucosidase does not act on the native polysaccharide but on the limit dextrin which remains when phosphorylase action is complete Other differences between the two enzymes are that

which is the same as that shown in Figure 2 The formation of phosphorylase *b* by the action of phosphorylase rupturing (PR) enzyme and the fact that *b* has half the molecular weight of *a* are accounted for as follows



$$\text{En UMP-PMU En}' + H_2O \rightarrow \text{En UDP} + \text{En}' \text{ UR} \quad (9)$$

(En UDP = phosphorylase *b*, UDP = uridine diphosphate, UR = uridylic acid)

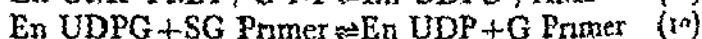
The participation of adenylic acid (AMP) in the action of phosphorylase *b* can be also accounted for



$$\text{En UDP} + \text{AMP} \rightleftharpoons \text{En UMP-PMA} + \text{P} \quad (10)$$



$$\text{En UMP PMA} + \text{G 1-P} \rightleftharpoons \text{En UDPG} + \text{AMP} \quad (11)$$



$$\text{En UDPG} + \text{SG Primer} \rightleftharpoons \text{En UDP} + \text{G Primer} \quad (12)$$

The sum of equations (10) (12) is the same as for phosphorylase *a*, i e equation (8) Other aspects of the relation of *a* to *b*, such as the conversion of *b* into *a* are accounted for in the same hypothesis It will be of interest to learn whether uridine diphosphate glucose also participates in the reactions of plant phosphorylases *Flavine adenine dinucleotide has been found in a jack bean phosphorylase preparation* (*Sumner*, Chou and Bever, 1950)

It is not clear to what extent, if any, phosphorylases *a* and *b* differ in their specific requirements towards the substrates for glycogen synthesis The investigations to be described have been made with phosphorylase *a* The fact that a primer is essential to the reaction was discovered in 1939 by Cori and Cori, who used glycogen for this purpose Linear α 1 4 linked glucose chains of the same length as the glycogen unit chains do not prime the reaction (Cori Swanson and Cori, 1945 Cori, 1954) an important distinction from potato phosphorylase (see above) *Glycogen amylopectin and amylose prime the* muscle phosphorylase reaction in proportion to their content of non reducing end groups but there is an upper limit of about 200 units on the length of chain capable of exerting priming activity (Cori *et al* 1945) The chain length of the polysaccharide present when synthesis from glycogen primer had ceased was determined by Hassid Cori and McCready (1953) In

Nagabhushanam, Nigam and Belavadi, 1955) If D enzyme were present in the amylo 1 6 glucosidase preparation it would not be necessary to postulate that the phosphorylase produced the glucose stubs, since these could be exposed by the trans ferring action of D enzyme

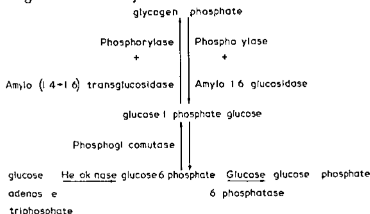


FIG 8 The enzymic synthesis and degradation of glycogen (Cori 1954)

The products of the combined or successive actions of phos phorylase and amylo 1 6 glucosidase are glucose 1 phosphate (from α 1 4 linked glucose units) and glucose from α 1 6 linked glucose units The molar ratio of these two products is there fore a measure of the average length of the unit chains in the parent polysaccharide (Cori and Larner, 1951, Illingworth, Larner and Cori 1952)

The interrelations of the glycogen metabolizing enzymes are shown in Figure 8 It has proved possible to relate some of the known glycogen storage diseases to the lack of sufficient amounts of one or other of these enzymes either by deductions based on the structure of the glycogen isolated from the diseased liver or by direct assay of the enzyme Storage diseases related to deficiencies in glucose 6 phosphatase amylo 1 6 glucosidase and amylo-(1 4→1 6) transglucosidase have been recognized (Illingworth and Cori 1952 Cori and Cori, 1952 Cori, 1954 Cori and Schulman 1954, Illingworth Cori and Cori 1956)

amylo-1 6 glucosidase also acts on amylopectin limit dextrin and hydrolyses isomaltose whereas R enzyme acts on amylopectin but not glycogen and does not attack isomaltose The action of amylo 1 6 glucosidase is recognized by the liberation of glucose and the susceptibility of the treated dextrin to further

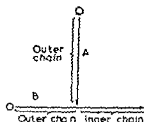


FIG 7 Types of unit chain in glycogen and amylopectin as defined by Peat Whelan and Thomas (1952) (O = non reducing-end group ⌇ = α - 1 6 link) Each molecule also contains one C chain which is a B chain with a free reducing-end group In the completely random form of the Meyer and Bernfeld (1940) model there are equal numbers of A and B (+C) chains (Hirst and Manners 1954) Only the outer chains are attacked by β amylase and phosphorylase α amylase can attack outer and inner chains

phosphorolysis It is believed that the glucose represents the residue of the A chains of the glycogen remaining after phosphorolysis (Figure 7) and therefore that phosphorylase can remove all but one of the glucose units in an A chain R Enzyme does not attack such terminally linked glucose units only non terminal 1 6 bonds (Whelan 1953) Phosphorolysis of the B chain (Figure 7) ceases when 5 or 6 glucose units remain on the non reducing side of the branch point (Cori and Larner, 1951) Since phosphorylase action is reversible there is a direct implication that these single glucose stubs the degraded A chains, are the priming nuclei for phosphorylase action but Cori and Larner (1951) have implied that this is not the case It does not seem possible that phosphorylase would so attenuate a glycogen chain that it lost its priming capacity It would be of interest to know whether rabbit muscle or liver contains any counterpart of D enzyme Such an enzyme capable of disproportionating maltose has been reported in rat liver and brain (Giri

Dextransucrase Dextransucrase from *Leuconostoc mesenteroides* and other bacteria, catalyses the conversion of sucrose into dex tran, a polymer of α 1 6 linked glucose units and free fructose (Hehre, 1941, see reviews by Hehre (1951) and Stacey and Ricketts (1951), Tsuchiya *et al* 1955 Barker Bourne Grant and Stacey 1956) Some dextrans are branched and it is pos sible that the dextransucrase preparations contain branching enzyme(s) as well as the chain forming catalyst As with amylosucrase action the conversion of sucrose is almost (99 per cent) complete, and it is difficult to decide whether sucrose is both the donor and acceptor substrate If it can act as an acceptor it is certainly very inefficient since the molecular weights of the polymers synthesized would require only 0 002 per cent of the sucrose to participate in this way (Edelman, 1956) This priming action might therefore be due to an impu rity in the sucrose or the enzyme Sugars such as glucose isomaltodextrins (α 1 6 linked glucose oligosaccharides) meli biose etc can function as acceptors since their presence in the reaction medium causes the formation of other oligosaccharides and a lowering in molecular weight of the polysaccharide which can be synthesized from a given amount of sucrose, i e the transferred glucosyl units are distributed between a greater member of acceptor molecules than in the absence of the test sugar (Koepsell *et al* 1953 Tsuchiya *et al* 1955)

Levansucrase An enzyme from *Bacillus subtilis* or *Aerobacter levanicum* transfers the β fructofuranosyl units of sucrose and builds a high molecular weight polymer of β 2 6 linked fruc tose residues (levan) liberating the glucose component as the free sugar (Hestrin Avineri Shapiro and Aschner, 1943, Hes trin Feingold and Avigad, 1956 Feingold Avigad and Hes trin 1956) The equilibrium lies far to the side of the polysac- charide and as with previous examples the nature of the primer from which the enzyme synthesizes high molecular weight levan is uncertain The enzyme can also transfer fructose to a variety of acceptor sugars forming sucrose analogues (Feingold Avigad and Hestrin, 1957)

Dextran dextrinase Aerobacter capsulatum and *viscosum* contain an enzyme capable of transforming short maltodextrin molecules

Polysaccharide Synthesis by Microbial Enzymes

Some of the micro-organisms which store starch or glycogen like polysaccharides seem to contain enzymes similar to those which plants and animals use for the synthesis of these substances For example, phosphorylase has been obtained from yeast (Meyer and Bernfeld, 1942, see Whelan, 1955), which also contains a debranching enzyme (Manners and Khin Maung, 1955) *Polytomella coeca* stores a starch containing amylose and amylopectin (Bourne, Stacey and Wilkinson, 1950), and is known to contain phosphorylase and Q enzyme (Bebbington, Bourne, Stacey and Wilkinson, 1952, Bebbington, Bourne and Wilkinson, 1952, Barker *et al*, 1953) Apart from these examples a few cases are known of polysaccharide syntheses which as yet have not found counterparts in the higher animals or plants These are described below

Amylosucrase This enzyme system has been detected in cell free extracts of *Neisseria perflava* (Hehre and Hamilton 1946, 1948 Hehre Hamilton and Carlson, 1949) It catalyses the conversion of sucrose into glycogen but it is probable that the product of amylosucrase action is amylose and that this is converted by a second enzyme (or enzymes) into glycogen, since it has been shown that this latter conversion takes place when the extract is incubated with amylose The fructose component of the sucrose is released as the free sugar Considerations of the energy associated with the sucrose and α 1 4 glucosidic linkages, the latter being much the stronger of the two, suggest that the equilibrium in the reaction will be to the side of the polysaccharide and in fact a 98 per cent conversion into glycogen was found The reversibility of the reaction was demonstrated by incubating starch with fructose and adding dextransucrase (see below) to take up the sucrose as it was formed and to turn it into dextran (Hehre and Hamilton 1951) It is not certain whether a special primer is needed by amylosucrase or whether sucrose itself will accept the α glucosyl units transferred from another sucrose molecules but it is known that synthesis does not take place in the presence of α amylase, which destroys plant phosphorylase primers (Hehre, 1951)

is the case with the maltodextrins in D enzyme action the trans ferred glycosyl radicals will be shared between all the sugar molecules present instead of a limited and controllable number of added primer (acceptor) molecules Two factors which influence the position of equilibrium in these reactions are the relative energy levels of the glycosidic linkages before and after transfer and the acceptor activity of the residue which remains after transfer from the donor, i e the molecule HOR in equation (1) A large energy difference will lead to a high conversion of the donor, e g the conversion of sucrose by dextransucrase The acceptor efficiency of HOR is determined by the activation energy required for its participation in a reaction with a donor, and the energy of the resultant linkage An example in which HOR has appreciable acceptor activity is the synthesis of amylose, in which inorganic phosphate (HOR) competes with the primer molecules (HOA) for glucosyl units transferred from glucose 1 phosphate, to the extent that the equilibrium quotient [inorganic phosphate]/[glucose 1 phosphate] is 3 2 at pH 7 0 The polysaccharide primer molecules are therefore appreciably more efficient as acceptors than is inorganic phosphate, but the difference in acceptor activity between amylose and inorganic phosphate cannot be nearly so great as between dextran molecules and fructose in the formation of dextran by dextransucrase

Chemical Synthesis of Polysaccharides

By comparison with the ease and specificity with which enzymes synthesize polysaccharides chemical methods are crude and cumbersome Monosaccharides can be polymerized by treatment with acid The main condition conducive to the formation of large molecules by this means is that the water content of the reaction mixture should be at a minimum Polysaccharides made in this way from glucose contain mainly α and β 1 6 linkages but any and every type of glucosidic linkage is likely to be present to some extent (Pacsu and Mora 1950 Kent 1953 Ricketts 1954 Thompson Anno Wolfrom and Inatome, 1954) A method for controlling the type of linkage formed has been devised by Haq and Whelan (1956a) who showed that 2 3 4

into a dextran like molecule of apparently high molecular weight (Hehre and Hamilton, 1949, 1951) The most active donor substrate tested was maltoheptaose Large amylaceous molecules (amylose, amylopectin and glycogen) are not glucose donors This is claimed as the first known instance of the conversion of one polymeric material into another in which essentially every linkage is affected' It would be of interest to know more about the acceptor requirements of this enzyme

Amylomaltase When *Escherichia coli* is adapted to grow on maltose there is developed an enzyme with the ability to transfer glucose units between maltose molecules, producing glucose and a mixture of low molecular weight maltodextrins (Monod and Torriani, 1950 Monod, 1950, Barker and Bourne, 1952) Like D enzyme this is not a true polysaccharide synthesizing enzyme but can be induced to make large molecules by a continuous removal of the glucose as was shown by adding glucose oxidase to the amylomaltase maltose system, when a blue iodine staining polysaccharide was formed (Monod and Torriani 1950)

ENZYMIC SYNTHESIS OF OLIGOSACCHARIDES

As mentioned earlier there are now a considerable number of examples of this type of synthesis These are outside the scope of this discussion and the reader is referred to the review by Edelman (1956) The manner in which oligosaccharide synthesis takes place is exactly the same as for polysaccharide synthesis i e by transfer of glycosyl radicals from a donor substrate to an acceptor substrate but it is instructive to explore the reasons why in some instances the products never attain a high degree of polymerization while in others the synthetic molecules may contain hundreds or thousands of monosaccharide residues The essential requirements for polysaccharide, as opposed to oligosaccharide synthesis seem to be first that the enzyme must not utilize water as an acceptor substrate for this would lead to the ultimate breakdown of the polymers as in the action of invertase on sucrose (Bacon 1953 Edelman, 1956) and secondly that the donor substrate must have no or very weak, acceptor activity If the acceptor activity is strong as

CHEMICAL DEGRADATION OF POLYSACCHARIDES

Methods for the chemical degradation of polysaccharides are legion and it is proposed to mention only those which resemble enzymic degradations These are a pointer to the possibility that we may one day be able exactly to imitate enzymic action with synthetic catalysts Acid hydrolysis corresponds to the random fragmenting action of an endo enzyme while an exo action can be obtained with alkali, periodate or by the Finan and O Colla degradation method, as explained below

Alkaline degradation This method has largely been developed by Kenner and co workers (Kenner and Richards, 1956, Bottle and Gilbert, 1954) and has been applied mainly to 1 4 and 1 3 linked glucose polymers An endwise splitting of monosaccharide residues takes place action commencing at the reducing chain end A 1 3 linked glucose unit is converted into *meta*saccharinic acid and a 1 4 linked unit into *iso*saccharinic acid There is an analogy with the exo action of β amylase in that alkaline degradation cannot pass through a 1 6 linkage

Periodate oxidation Periodate salts or the free acid oxidize α glycol groupings and therefore can effect a considerable oxidation of polysaccharide molecules (see review by Bobbitt (1956)) After the initial oxidation a slower reaction may set in whereby the molecule is eroded in stepwise manner from the reducing end Chains of 1 3 or 1 4 linked hexose units are attacked but 1 6 linked molecules do not undergo erosion and a 1 6 link in the interior of a 1 4 linked chain is a barrier to further oxidation (Lawley 1955, Bines 1956 Hough and Perry 1956 Parrish unpublished results)

Barry (1943) degradation When a periodate oxidized polysaccharide is heated with phenylhydrazine the oxidized portions of the molecule are broken off and the residue may become susceptible to further oxidation and degradation In the case of branched polysaccharides the successive degradations proceed essentially from the non reducing ends into the interior of the molecule (Dillon O Ceallachain and O Colla, 1953 Barry and Mitchell 1954)

Finan and O Colla degradation When the phenylosazone of

tri O acetyl α D glucopyranosyl bromide will undergo self condensation, with elimination of hydrogen bromide, to form chains of β 1 6 linked glucose units (gentiodextrins) Molecules containing up to 9 glucose units were detected A glucose anhydride has also been caused to polymerize in this way (Haq and Whelan, 1956b)

Enzymic Degradation of Polysaccharides

If by the term degradation we include any process in which sugar to-sugar linkages are decreased in number then the phosphorolysis of starch or glycogen to glucose 1 phosphate is a degradative process That is, a degradative process may be regarded as one in which the acceptor molecule is not a sugar The most ubiquitous of such acceptors is water and the enzymes which use water are termed hydrolases There are two main groups of carbohydrate hydrolases the *endo* and the *exo* hydrolases An *endo* enzyme is one which can attack linkages throughout the polysaccharide at random, breaking the molecule into smaller and smaller fragments An *exo* enzyme is restricted in its action to the end of the chain molecule and in all known cases it is the non reducing chain end at which reaction commences Typical of the endo enzymes is α amylase, found in almost all living organisms This hydrolyses the α 1 4 glucosidic linkage of amylaceous substances and in general prefers to attack linkages not situated at chain ends It does not attack the α 1 6 branch linkages β Amylase, found only in plants, is an exo enzyme, acting on chains of α 1 4 linked glucose units by splitting the penultimate and subsequent alternate linkages, commencing at the non reducing chain ends The disaccharide maltose is therefore the product of its action Other exo enzymes split only one monosaccharide at a time from the polysaccharide, e g glucamylase (Phillips and Caldwell 1951) The known examples of carbohydrate hydrolase action are far more numerous than of synthesis mainly because degradation is so much more easy to detect than synthesis The reader is referred to Manners (1955) and Whelan (1957) for detailed reviews of these enzymes

in oviduct (Strominger, 1955), neuraminic and sialic acid like compounds (Park 1952) glucose, galactose glucuronic acid *N* acetylglucosamine, arabinose and xylose, in mung bean seedlings (Ginsberg Stumpf and Hassid 1956 Sohns, Feingold and Hassid 1957)

All the known examples of polysaccharide synthesis involve the transfer of single monosaccharide residues from the donor to the acceptor substrate But this may not prove to be the rule, and in the search for donor substrates consideration should be given to the possibility that the building units may be oligo- or poly saccharides Such a mode of synthesis would rationalize the enzymic requirements for the synthesis of the complex poly saccharides Q Enzyme action is in effect an example of this type of synthesis The suggestion may be illustrated by refer ence to particular polysaccharides Nigeran, from *Aspergillus niger*, is a glucose polymer in which 1 3 and 1 4 links alternate (Barker, Bourne and Stacey 1953 Barker, Bourne O Mant and Stacey, 1957) It was suggested by Barker and Carrington (1953) that nigeran may not be synthesized by the addition of glucose units in alternating linkage to a primer but by the union of maltose radicals (α 1 4 linked) through α 1 3 links or by join ing radicals of the α 1 3 linked diglucose (nigerose) through α 1 4 bonds Similarly Peat Whelan and Roberts (1957) have suggested that lichenin (from Iceland moss) may be synthesized by polymerization of cellotriose radicals through α 1 3 links since the linkages in this polysaccharide seem to have the repeat ing pattern β 1 4 β 1 4 β 1 3 A barley gum (Aspinall and Telfer 1954) is another possible instance of a regular repeating structure When a more detailed examination has been made of the polysaccharides containing several different monosac charides e g the blood group polysaccharides or the plant gums it may be found that they are based on repeating groups of sugars It may also be that this type of synthesis is not re stricted to complex polysaccharides If lichenin is indeed built up from cellotriose radicals joined through β 1 3 links may not cellulose be synthesized by joining cellotriose through β 1 4 links? This type of synthesis would moreover indicate the *in vivo* function of the oligosaccharide synthesizing enzymes These

laminaritriose (β 1 3 linked glucose trisaccharide) was treated with a cation exchange resin at room temperature there were produced laminaribiose and glucosazone (Finan and O Colla 1955) Osazone formation takes place at the reducing groups of sugar molecules so that there is provided a method for the stepwise degradation of reducing sugar polymers from their reducing chain ends

Conclusion

One of the most pressing problems in studies of polysaccharides is the question of the mode of their enzymic synthesis It is a sobering thought as already mentioned that the main outlines of the enzymic synthesis of glycogen and starch were established in the years 1937-44 and that these are still the only instances of polysaccharide synthesis in plants and animals where the responsible enzymes have been isolated and the *in vitro* synthesis is possible The problem at the moment is not so much to obtain the polysaccharide synthesizing enzymes in the active state as to discover the donor and acceptor substrates on which they act The fortunate chance that phosphorylase action is reversible was responsible for the discovery of glucose 1 phosphate but many polysaccharide syntheses will probably prove to be practically irreversible as with syntheses based on sucrose as substrate (see above) so that the search for the substrate must be largely empirical A systematic examination of the particular animal or plant part which is the site of polysaccharide synthesis for possible substrates would seem to be the only way of systematizing the search It may be that uridine phosphate sugar compounds will prove to be key intermediates [1] The role of uridine diphosphate glucose in sucrose and glycogen synthesis has already been mentioned Uridine diphosphate derivatives of the following sugars have also been discovered glucuronic acid in liver (Storey and Dutton 1955, Strominger, Maxwell, Axelrod and Kalckar 1957) *N* acetylglucosamine and *N* acetylgalactosamine the former in yeast (Cabib Leloir and Cardini 1953) and both in liver (Pontis, 1955) *N* acetyl galactosamine sulphate and *N* acetylglucosamine 6 phosphate

[1] See p 233 n 1

CARDINI C E LELOIR L F and CHIRIBOGA J (1955) *J biol Chem* **214** 149
COHN M (1949) *J biol Chem* **180** 771
CORI G T (1954) *Harvey Lect* Ser 48 145
CORI G T (1955) *Methods in Enzymology* vol 1 Academic Press New York.
CORI G T and CORI C F (1939) *J biol Chem* **131** 397
CORI G T and CORI C F (1943) *J biol Chem* **151** 57
CORI G T and CORI C F (1952) *J biol Chem* **199** 661
CORI G T ILLINGWORTH B and KELLER P J (1955) *Methods in Enzymology* vol 1 Academic Press New York
CORI G T and LARNER J (1951) *J biol Chem* **188** 17
CORI G T and SCHULMAN J L (1954) *Pediatrics Springfield* **14** 632
CORI G T SWANSON M A and CORI C F (1945) *Fed Proc* **4** 234
DILLON T O CEALLACHAIN D F and O COLLA P A (1953) *Proc R Irish Acad* **55B** 331
DOUDOROFF M BARKER H A and HASSID W Z (1947) *J biol Chem* **168** 725 733
EDELMAN J (1956) *Advanc Enzymol* **17** 189
FEINGOLD D S AVIGAD G and HESTRIN S (1956) *Biochem J* **64** 351
FEINGOLD D S AVIGAD G and HESTRIN S (1957) *J biol Chem* **224** 295
FINAN P A and O COLLA P S (1955) *Chem & Ind* p 1387
FRENCH D and WILD G M (1953) *J Amer chem Soc* **75** 4990
GINSBERG V STUMPF P K and HASSID W Z (1956) *J biol Chem* **223** 977
GIRI K V NAGABHUSHANAM A NIGAM V N and BELAVADI B (1955) *Science* **121** 898
HANES C S (1940a) *Proc Roy Soc* B **128** 421
HANES C S (1940b) *Proc Roy Soc* B **129** 174
HAQ S and WHELAN W J (1956a) *J chem Soc* p 4543
HAQ S and WHELAN W J (1956b) *Nature Lond* **178** 122
HASSID W Z CORI G T and MCCREADY R M (1943) *J biol Chem* **148** 89
HAWORTH W N PEAT S and BOURNE E J (1944) *Nature Lond* **154** 236
HEHRE E J (1941) *Science* **93** 237
HEHRE E J (1951) *Advanc Enzymol* **11** 297
HEHRE E J and HAMILTON D M (1946) *J biol Chem* **166** 777
HEHRE E J and HAMILTON D M (1948) *J Bact* **55** 197
HEHRE E J and HAMILTON D M (1949) *Proc Soc exp Biol N Y* **71** 336
HEHRE E J and HAMILTON D M (1951) *J biol Chem* **192** 161
HEHRE E J HAMILTON D M and CARLSON A S (1949) *J biol Chem* **177** 267
HESTRIN S (1949) *J biol Chem* **179** 943

could be providing the building units for polysaccharide synthesis

REFERENCES

ABDEL AKHER, M and SMITH F (1951) *J Amer chem Soc* 73 994
ASPINALL G O and TELFER R G J (1954) *J chem Soc* p 3519
BACON J S D (1953) *Rep Progr Chem* 50 281
BARKER S A BEBBINGTON A and BOURNE E J (1953) *J chem Soc* p 4051
BARKER S A and BOURNE E J (1952) *J chem Soc* p 209
BARKER S A and BOURNE E J (1953) *Quart Rev chem Soc Lond* 7 56
BARKER S A BOURNE, E J GRANT P M and STACEY M (1956) *Nature Lond* 178 1221 also (1957) *J chem Soc* pp 3530 3536
BARKER S A BOURNE E J O MANT D M and STACEY, M (1957) *J chem Soc* p 2448
BARKER S A, BOURNE E J PEAT S and WILKINSON I A (1950) *J chem Soc* p 3022
BARKER S A, BOURNE E J and STACEY, M (1953) *J chem Soc* p 3084
BARKER, S A BOURNE E J and WILKINSON I A (1950) *J chem Soc* p 3027 see also earlier papers in this series
BARKER S A and CARRINGTON T R (1953) *J chem Soc* p 3588
BARRY V C (1943) *Nature Lond* 152 537
BARRY V C and MITCHELL P W D (1954) *J chem Soc* p 4020
BAUM H and GILBERT G A (1953) *Nature Lond* 171 983
BAUM H GILBERT G A and SCOTT N D (1956) *Nature Lond* 177 889
BEAN R C and HASSID W Z (1955) *J Amer chem Soc* 77 5737
BEBBINGTON, A BOURNE E J STACEY M and WILKINSON I A (1952) *J chem Soc* p 240
BEBBINGTON A, BOURNE E J and WILKINSON I A (1952) *J chem Soc* p 246
BINES, B J (1956) Ph D thesis University of Wales
BOBBITT J M (1956) *Advanc Carbohyd Chem* 11 1
BOTTLE R T and GILBERT G A (1954) *Chem & Ind* p 1201
BOURNE E J SITCH D A and PEAT S (1949) *J chem Soc* p 1448
BOURNE E J STACEY M and WILKINSON I A (1950) *J chem Soc* p 2694
BUCHANAN J G (1953) *Arch Biochem Biophys* 44 140
BUNTON C A LEWIS T A LLEWELLYN D R TRISTAM H and VERNON C A (1954) *Nature Lond* 174 560
BURMA D P and MORTIMER D C (1956) *Arch Biochem Biophys* 62 16
CABIB, E LELOIR L F and CARDINI C E (1953) *J biol Chem* 203 1055
CALVIN M (1956) *J chem Soc* p 1895

PEAT S WHELAN W J and BAILEY J M (1952) *J chem Soc* p 1422
PEAT S WHELAN W J HOBSON P N and THOMAS G J (1954) *J chem Soc* p 4440
PEAT S WHELAN W J and JONES G (1957) *J chem Soc* p 2490
PEAT S WHELAN W J and REES W R (1956) *J chem Soc* p 44
PEAT S WHELAN W J and ROBERTS J G (1957) *J chem Soc* p 3916
PEAT S WHELAN W J and THOMAS G J (1952) *J chem Soc* p 4546
PEAT S WHELAN W J and THOMAS G J (1956) *J chem Soc* p 3025
PHILLIPS L L and CALDWELL M L (1951) *J Amer chem Soc* **73** 3559 3563
PONTIS H G (1955) *J biol Chem* **216** 195
RICKETTS C R (1954) *J chem Soc* p 4031
SOHNS J FEINGOLD D S and HASSID W Z (1957) *J Amer chem Soc* **79** 2342
STACEY M and RICKETTS C R (1951) *Fortschr Chem org Naturst* **8** 28
STOREY I D E and DUTTON G J (1955) *Biochem J* **59** 279
STROMINGER J L (1955) *Biochim biophys acta* **17** 283
STROMINGER J L MAXWELL E S AXELROD J and KALCKAR H M (1957) *J biol Chem* **224** 79
SUMNER J B CHOU T C and BEVER A. T (1950) *Arch Biochem Biophys* **26** 1
SUTHERLAND E W (1955) *Methods in Enzymology* vol 1 Academic Press New York
SWANSON M A. and CORI C F (1948) *J biol Chem* **172** 815
THOMPSON A ANNO K WOLFROM M L and INATOME M (1954) *J Amer chem Soc* **76** 1309
TSUCHIYA H M HELLMAN N N KOEPSELL H J CORMAN J STRINGER C S ROGOVIN S P BOGARD M O BRYANT G FEGER V H, HOFFMAN C A SENTI F R. and JACKSON R W (1955) *J Amer chem Soc* **77**, 2412
WALKER G J and WHELAN W J (1957) *Biochem J* **67** 548
WEIBULL C and TISELIUS A (1945) *Arkiv Kemi Min Geol* A19 no 19
WHELAN W J (1953) *Biochem Soc Symposia* **11** 17
WHELAN W J (1955) *Methods in Enzymology* vol 1 Academic Press New York
WHELAN W J (1957) *Hanbuch der Pflanzenphysiologie* vol VI Springer Verlag Berlin
WHELAN W J and BAILEY J M (1954) *Biochem J* **58** 560
WHISTLER R L and SMART C L (1953) *Polysaccharide Chemistry* Academic Press New York
WOLFROM M L and THOMPSON A (1956) *J Amer chem Soc* **78** 4116

HESTRIN, S, AVINERI SHAPIRO S and ASCHNER M (1943) *Biochem J* **37** 450
HESTRIN S FEINGOLD D S and AVIGAD G (1956) *Biochem J* **64** 340
HIRST E L and MANNERS D J (1954) *Chem & Ind* p 224
HOBSON P N WHELAN W J and PEAT S (1951) *J chem Soc* p 1451
HOUGH L and PERRY M B (1956) *Chem & Ind* p 768
ILLINGWORTH B and CORI G T (1952) *J biol Chem* **199** 653
ILLINGWORTH B CORI G T and CORI, C F (1956) *J biol Chem* **218** 123
ILLINGWORTH B LARNER J and CORI G T (1952) *J biol Chem* **199** 631
JERMYN M A (1957) *Science* **125** 12
KENNER J and RICHARDS G N (1956) *J chem Soc* p 2921 and earlier parts of this series
KENT, P W (1953) *Biochem J* **55** 361
KOEPSELL H J TSUCHIYA H M, HELLMAN N N KAZENKO A HOFFMAN C A SHARPE E S and JACKSON R W (1953) *J biol Chem* **200** 793
KORKES S (1956) *Ann Rev Biochem* **25** 685
KOSHLAND D E and STEIN S S (1954) *J biol Chem* **208** 139
LARNER J (1953) *J biol Chem* **202** 491
LARNER J (1955) *Methods in Enzymology* vol 1 Academic Press New York
LARNER J ILLINGWORTH B CORI G T and CORI C F (1952) *J biol Chem* **199** 641
LARNER J and SCHLISELFELD L H (1956) *Biochim biophys acta* **20** 53
LARNER J and UWAH D N (1956) *J Amer chem Soc* **78** 3647
LAWLEY H G (1955) Ph D thesis University of Wales
LELOIR L F and CARDINI C E (1955) *J biol Chem* **214** 157
LINKER A MEYER K and HOFFMAN P (1956) *J biol Chem* **219** 13
MANNERS D J (1953) *Rep Progr Chem* **50** 288
MANNERS D J (1955) *Quart Rev chem Soc Lond* **9** 73
MANNERS D J and ARCHIBALD A R (1957) *J chem Soc* p 2205
MANNERS D J and KHIN MAUNG (1955) *Chem & Ind* p 950
MEYER K H and BERNFELD P (1940) *Helv chim acta* **23** 875
MEYER K H and BERNFELD P (1942) *Helv chim acta* **25** 399
MONOD J (1950) *Biochem Soc Symposia* **4** 51
MONOD J and TORRIANI A M (1950) *Ann Inst Pasteur* **78** 65
NUSSENBAUM S and HASSID W Z (1952) *J biol Chem* **196** 785 *Analyt Chem* **24** 501
PACSU E and MORA P T (1950) *J Amer chem Soc* **72** 1045
PARK J T (1952) *J biol Chem* **194** 885
PAZUR, J H BUDOVICH T and TIPTON C L (1957) *J Amer chem Soc* **79** 625
PEAT S THOMAS G J and WHELAN W J (1952) *J chem Soc* p 722
PEAT S TURVEY J R and EVANS J M (1957) *Nature Lond* **179** 261

It seems that vitamin B_6 (pyridoxin) is required for the conversion of linoleic to arachidonic acid (Witten and Holman, 1952)

There are certain chemical points that are important Obviously a large number of isomers of linoleic acid can exist either positional isomers in which the double bonds are shifted along the molecule to different positions or geometrical isomers in which the hydrogens on one or both double bonds are *trans* (giving *cis trans*, *trans cis* or *trans trans* isomers—hereafter referred to collectively as *trans*) Unfortunately the *trans* isomers are more stable than the natural *cis cis* isomer It will be noticed that the double bonds are methylene interrupted and not in the more stable conjugated position (–CH=CH–CH=CH–) and that in each of the four compounds shown there is no double bond until the sixth carbon from the methyl end is reached, if a double bond occurs in this part (as in the ordinary form of linolenic acid) the compound has no EFA activity and this is important since many highly unsaturated (or polyethenoid) fatty acids present in fish oils have a double bond in the 3 4 position counting from the methyl end

The presence of double bonds makes EFA easily oxidized if oxygen is present and traces of metals catalyse this oxidation with production of hydroperoxides which are themselves strong oxidizing agents Reduction also occurs fairly easily under appropriate conditions with production of stearic acid nickel is used for catalysing this hydrogenation in the manufacture of margarine and if hydrogenation is not complete *trans* and positional isomers of linoleic acid are formed So unstable are the EFA in presence of oxygen that they would not exist for long in the food or body were it not for the presence of antioxidants the most important of which are the tocopherols (vitamins E) These are distributed in plant sources roughly parallel with the distribution of linoleic acid They hinder destruction of EFA in the food in the gut and in the body There appears to be some synthesis of EFA in the gut by bacteria and in the body when we speak of arachidonic acid and its precursors as being 'essential fatty acids we mean that one or other or combinations must be supplied in the diet since they cannot be synthesized in the

XV

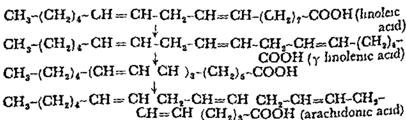
Essential Fatty Acids in Nutrition and their relation to other Vitamins

H M SINCLAIR

Nature of Essential Fatty Acids

The need for fat in the diet of animals was not demonstrated until Evans and Burr (1928) showed that a diet of purified casein and sucrose, supplemented with salts and yeast and cod liver oil, produced subnormal growth in rats Various workers have now studied the syndrome of deficiency of essential fatty acids (EFA) in at least seven species of animal including monkeys, and there are strong reasons for believing that they are needed by man

The most important EFA in the diet is linoleic acid or *cis* 9 10, *cis* 12 13 octadecadienoic acid, this has one more double bond than has oleic acid and two more than stearic acid all of which have 18 carbon atoms Linoleic acid is converted in the body to γ linolenic acid by the insertion of a third double bond but this acid does not occur in appreciable amounts in food stuffs Then two carbons are added to γ linolenic acid and a fourth double bond is inserted giving arachidonic acid which is all *cis* 5 6 8 9 11 12, 14 15 eicosatetraenoic acid These steps which have been elucidated by Mead and Howton (1957) are as follows



The presence of these isomers may be important The requirement of EFA in the diet depends upon various factors one of the most important of which is the proportion of EFA to other fatty acids Saturated fat, such as stearic acid increases the requirement and there is some evidence that isomers of linoleic acid compete with the ordinary form in the body and so tend to produce deficiency on marginal diets This is a matter of great importance about which we have very little information I mentioned it after we had done some work upon the presence of isomers in foodstuffs (Sinclair, 1956a) but we stopped our work on this when the elegant and extensive work of J B Brown appeared He has shown (Mabrouk and Brown, 1956 Sreenivasan and Brown 1956) that typical margarines may contain about 40 per cent of *trans* acids and that about half the linoleic acid may be conjugated (and therefore not effective as EFA)

Wheat germ oil is rich in linoleic acid and in tocopherols and more of the former than one might expect is found in white flour since the starch adsorbs linoleic acid However, as I have recently pointed out elsewhere (Sinclair 1957c d), it is known from the work of various authors that most of the tocopherols is destroyed in flour by the bleaching improvers such as chlorine dioxide, and there is some evidence that this destroys and chlorinates some of the linoleic acid though whether this is significant is not at present known We have to consider both the composition of the fat in foodstuffs and the amount of that fat ingested and unfortunately we have not adequate analyses at present the fat of potatoes is about 50 per cent linoleic acid but there is so little fat present that this is not important

We are even more ignorant about the contribution of animal fats and their composition is changing in some instances Some fish oils are very rich in EFA, but some marine oils, such as cod liver oil are extremely deficient and contain polyethenoid acids that are antagonistic cod liver oil is therefore toxic for EFA deficient rats or mice Eggs from hens free ranging and therefore ingesting linoleic acid may contain considerable quantities of linoleic and arachidonic acids but if the hens are kept in batteries and fed on poultry meal deficient in linoleic acid the eggs will also be deficient Lard used to be rich even in arachidonic

body in adequate amount Unfortunately bacteria also destroy EFA, and this occurs in the gut of ruminants such as cows and sheep, where bacteria convert *cis cis* linoleic acid from ingested plant fodder into *trans* isomers which are then stored in the body fat and excreted in milk

Sources of EFA

The richest foodstuffs are the vegetable seed oils, more than half of the fatty acids of some of them consisting of linoleic acid For instance the following are the approximate percentages of linoleic acid in the fatty acids of certain typical seed oils walnut, 73 safflower, 73, sunflower 60, corn, 55, soybean 50, cotton, 45 arachis 26, wheat 25 olive, 15 coconut, 2 Most of these oils are also rich in tocopherols and it is possible that the content of different tocopherols affects the ability of oils to lower serum cholesterol in man The tocopherols in corn oil consist of about 89 per cent of the γ and 11 per cent of α tocopherol The former is the more powerful antioxidant and therefore the better protects linoleic acid in the oil from oxidation in the air or in the gut but it is itself hardly absorbed from the gut and therefore, unlike the α form, is not effective in the body The combination of a large proportion of the very active unabsorbable form with a small proportion of the less active easily absorbed form is probably excellent Unfortunately we eat very little unprocessed vegetable seed oils in this country Refined olive oil is not a good source of EFA and the large quantities of vegetable seed oils that are imported are used for the manufacture of margarine Margarines vary considerably in their content of EFA it seems that in general French margarine is rich British margarine less rich and American margarine poor Part of the vegetable oil may be totally hydrogenated and part not hydrogenated at all, in which case the resulting margarine will be rich in EFA Alternatively all the vegetable seed oil may be partially hydrogenated in which case most of the *cis cis* linoleic acid may be lost since the nickel catalyst will isomerize it before hydrogenation occurs But British margarine is probably richer in EFA than is butter since as already indicated almost all the linoleic acid in this is *trans* isomers

then in combination with protein form lipoprotein It appears that EFA are needed for proper metabolism of cholesterol and of phospholipids and the need for essential fatty acids for these purposes may be regarded as part of their structural function Rather different however, is the possible need for the esterifica tion of vitamin D which will be mentioned later

There is as yet no direct evidence that EFA are needed for mitochondrial membranes but unsaturated fats occur in them Uncoupling of oxidative phosphorylation is found in deficiency of EFA (Klein and Johnson, 1954) and this might be caused by the defective structure of mitochondria Ramalingaswami and I found (1953) that there was defective structure of cartilage and surrounding mesenchymal ground substance in the ear of the rat deficient in EFA and there seems to be a general defect in con nective tissue throughout the body The possible relationship of this to disorders of the skin and of bone I have discussed else-where (1957b e) The greatly increased permeability of the skin to water (Sinclair 1952, Basnayake and Sinclair 1954 1955) is explicable in terms of the structural defect and the greatly in creased fragility and permeability of capillaries found in defi ciency of EFA (Kramar and Levine 1953) probably occurs through a defect in the connective tissue that supports the endo thelial cells of the capillaries Proteinuria and haematuria are sometimes found in such animals and appear to be caused by similar defects in the glomeruli of kidney

Function of Essential Fatty Acids in Man

Hansen and his colleagues (1957) have studied extensively defi ciency of EFA in children particularly in relation to infantile eczema They have found that in certain cases of this condition the iodine value of serum lipids is decreased as compared with normal infants and that treatment with fats rich in EFA cures some cases of eczema A good deal of supporting and also con flicting evidence has been produced in various quarters and I have recently reviewed it elsewhere (1957b) The balance of evidence is in favour of EFA causing at least some cases of eczema in infants

Several years ago we considered that some cases of follicular

acid when pigs ranging freely were fed on scraps and their fat was sold by a butcher, modern pigs, kept in Danish houses, are fed on processed meal and a premium is given for 'hard' fat (which means saturated), and then the fat is processed so that it will have a long shelf life in a grocer's shop Little is quantitatively known about the effect of cooking on the EFA in foods, but frying undoubtedly causes losses Therefore as we process our foods more, we tend to decrease the content of EFA and increase that of their isomers and of saturated fat Further, a diet excessively high in carbohydrate will give rise in the body to relatively saturated fat which will also increase the requirement of EFA in the diet Conversely, on very low calorie diets the body will burn EFA ingested and will mobilize the relatively saturated fat from the body stores so that EFA deficiency will tend to occur And there are other factors that affect the requirement it is increased in hypothyroidism and particularly in diabetes mellitus, in the latter case the explanation may be that EFA are swept into the general pool of fat that is rapidly being metabolized

These dietary considerations have not been adequately considered by those who have concluded that the relative EFA content of our diets has not decreased over the years (Hollingsworth, Vaughan and Warnock, 1956, McCann and Trulson, 1957 Yudkin, 1957) My own studies have led to the opposite conclusion and will be presented elsewhere I believe that during the past forty years there has been an increasing relative deficiency of EFA in our diets, although there have been fluctuations in certain years The epidemiology of deficiency of EFA in different countries of the world will be mentioned presently

Function of Essential Fatty Acids in Lower Animals

The main functions appear to be structural and with certain qualifications all the functions at present known can be explained in these terms EFA appear to be needed for the formation of cell membranes (including the myelin sheath of nerves), for mitochondrial membranes and the cristae mitochondriales, and for the formation of connective tissue In cell membranes EFA esterify cholesterol and form part of phospholipids, and

body, cholesterol esters of quite different physical properties are formed they have a higher melting point and are less soluble in organic solvents Similarly cholesterol esterified with the ordinary form of linoleic acid is more soluble and probably more easily metabolized in the body than that formed with conjugated linoleic acid By analogy with this deposition of cholesterol and phospholipid in the epidermis it seemed possible that the accumulation in other tissues might occur through the same mechanism (atheroma, arcus senilis cataract xanthoma tendinosum) This local synthesis and accumulation within the cells of the intima however, would not account for the patchy distribution of atheroma though it would account for the earliest lesion being found within the cells and not extracellularly as might be expected if the cholesterol diffused from serum As mentioned above there is increased permeability of capillaries in deficiency of EFA and this might allow diffusion from plasma to occur predominantly in areas where blood pressure is raised or where there is local trauma (as for instance near branches from the aorta)

The greatly increased requirement of EFA in diabetes mellitus has already been mentioned and the deposition of cholesterol in this condition is perhaps therefore not surprising It is tempting to explain the three other main complications of diabetes mellitus on a similar basis Lower animals deficient in EFA sometimes develop nephropathy for a reason that has already been mentioned and the defect in the connective tissue supporting the endothelial cells of venules might be responsible for the aneurysmal dilatations that occur in the retina in diabetic retinitis The nervous system is very rich in EFA but under normal conditions these are avidly retained when other tissues become depleted In long standing deficiency however or in very severe deficiency in lower animals depletion of the nervous system can occur and diabetic neuropathy might be caused in this way

One interesting fact about EFA is that the requirement for male animals is much greater than that of females during the reproductive period the ratio is about seven In some way not understood this seems to be related to the sex hormones There

hyperkeratosis in man might be caused by deficiency of EFA and we therefore carefully studied the morphology of the lesion of the skin produced by this deficiency in rats (Ramalingaswami and Sinclair, 1953) Our preliminary studies of certain cases of this condition and of Darier s disease and of pityriasis rubra pilaris have provided some support for this suggestion, though extensive clinical trials require to be done

The possible relation of EFA to certain diseases in man, which I have suggested speculatively elsewhere (Sinclair 1956a), can be most conveniently considered in the following sections

ATHEROMA

Apart from the liver there are seven tissues of the body in which cholesterol may be deposited and all of which have in common the anatomical peculiarity of not being supplied directly with capillaries or drained by lymphatics These avascular tissues are epithelium, endothelium cornea (both epithelium and endothelium) lens, tendon cartilage, the granulosa cells of the ovary and the cells of the seminiferous tubules in the testis In all these tissues deposits of cholesterol may occur for instance in diabetes mellitus or idiopathic hypercholesterolaemia These tissues (or their appendages such as nails and hair) together with teeth comprise the avascular tissues of the body In lower animals such as the rat it is not easy to study the deposition of cholesterol and phospholipid in the intima of the aorta but in the belief that the process might be similar in the epidermis we have studied this Both the intima and the epidermis are known to be able to synthesize cholesterol and to esterify it We found (Basnayake and Sinclair 1955) that the rat fed a diet totally deficient in fat (and therefore containing no EFA) deposited large amounts of cholesterol and phospholipid in its epidermis and that these were abnormal in composition in that the iodine number of the fatty acids was abnormally low The probable explanation for the deposition is that under normal conditions these compounds are formed with EFA and are then easily combined with protein to form lipoprotein and metabolized in the body If more saturated fatty acids are used such as the fat that can be synthesized in the

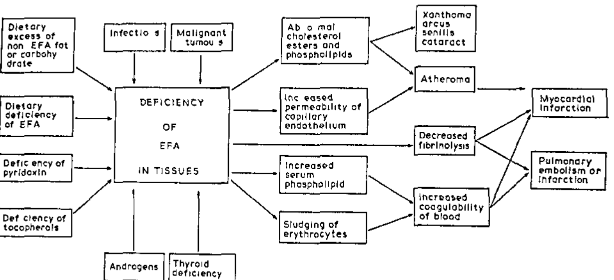


FIG 1 Diagram of aetiological factors in myocardial infarction

is also a possible relation of EFA with corticosteroids, as will be mentioned presently, and it may be that steroid hormones become esterified with EFA as do cholesterol and possibly vitamin D It seems that oestrogens, unlike androgens, increase the excretion in the bile of cholesterol and its breakdown products such as cholic acid Herein may lie the origin of gallstones which are prevalent in EFA deficient lower animals Gallstones occur in the more highly civilized countries, in women rather than men during the reproductive period, and in women who are obese and therefore more likely to be deficient in EFA In rats deficient in EFA there is desquamation of the epithelium of the gall bladder and thereby a nidus is provided for the formation of a stone The same may be true of man, and therefore there might be an increasing tendency for stones in women up to the menopause since oestrogens may increase the concentration of cholesterol in the gall bladder and therefore the tendency for stone to form on the desquamated epithelium

In a very different category are stones of the urinary bladder which are prevalent in undernourished boys in certain Eastern countries (Passmore, 1953) Here we may have EFA deficiency causing desquamation of the epithelium of the bladder and so providing a nidus, the prevalence in males might arise from the increased EFA requirement of males

Coronary Thrombosis

I have attempted elsewhere (Sinclair, 1956a 1957a) to put forward a unified hypothesis for coronary thrombosis on the basis of EFA deficiency Since then this has been produced experimentally in rats (Hartroft and Thomas, 1957) I have supposed as have others that atheroma and increased coagulability of the blood are required for myocardial infarction (Figure 1) Atheroma we have already considered, and it is interesting, as I shall show in detail elsewhere that deaths from myocardial infarction are increasing at about the same rate as are those ascribed to pulmonary embolism or infarction This apparent increase in coagulability of the blood of the population might be related to increasing relative deficiency of EFA by two factors First, Robinson and Poole (1956) have shown that phosphatidyl

of cholesterol Our own work on the deposition of cholesterol suggests that this is a function of EFA in particular rather than of polyethenoid acids in general though this is not proved EFA will remove the accumulated cholesterol and phospholipid from an EFA deficient rat or mouse but an attempt to do so with cod liver oil (rich in polyethenoid acids but deficient in EFA) caused death of the animals

Keys on the other hand maintained that all fats were equal in their cholesterol raising effects on plasma, though his more recent experiences with corn oil are in accord with those obtained earlier by other workers and cannot be reconciled with his own earlier results (see Kinsell and Sinclair, 1957) He has maintained that deaths from myocardial infarction in different countries run parallel with total fat in the dietary (1953), but only six countries were selected and some of the obvious misfits such as France were not included Yerushalmy and Hilleboe (1957) have correctly discounted his views I have used official figures to correlate deaths from myocardial infarction with assessed relative deficiency of EFA in dietaries and there is a reasonable correlation as will be presented elsewhere

Idiopathic Hypercalcaemia of Infants

This condition was first described by Lightwood in 1952 and since then many cases of the severe and mild types have been reported The infants almost invariably fed on dried cow s milk with added vitamin D, develop athrepsia anorexia vomiting polydipsia hypertension with a systolic murmur over the heart patchy osteosclerosis mental retardation and nephrocalcinosis, in serum calcium urea cholesterol (mainly esterified) and vitamin A are increased Bonham Carter (Bonham Carter *et al*, 1955) suggested that the systolic murmur was associated with hypertension and atheroma which accompanies this condition in some of its phases Some cases follow infection

It seems to be generally agreed that the condition is caused by increased sensitivity to vitamin D but this cannot explain all the features Elsewhere (Sinclair 1956b c) I have put forward a possible explanation supported by some experimental evidence on lower animals Cow s milk is known to be very deficient in

ethanolamine appears to be increased in plasma after a fatty meal and that it increases the coagulability of blood This type of phospholipid—the cephalin type—characteristically contains more highly unsaturated fatty acids than the lecithin type When Basnayake and I found (1955) that phospholipid accumulated *in the epidermis in EFA deficiency and subsequently that* the iodine number of the fatty acid in it was abnormally low, we considered that the accumulation was caused by failure to metabolize the abnormal compound Phospholipid is increased in plasma in myocardial infarction, and the origin of the increase might also be through failure to metabolize adequately an abnormal phospholipid that nevertheless has undiminished coagulating power EFA are also required for fibrinolysis'

The second fact of interest is that several years ago Ramalingaswami and I noticed when measuring the blood pressure of rats deficient in EFA or pyridoxin that, when the circulation in the limb was slowed by a cuff and the capillaries in the paw were observed, there was marked sludging of the erythrocytes in either deficiency but not in the controls This is similar to the phenomenon observed by Knisely (1951) in man, and might be expected to favour coagulation of blood

Various authors, particularly Kinsell *et al* (1957), Ahrens *et al* (1957), Malmros (Malmros and Wigand, 1957), and Bronte Stewart (1957), have shown that essential fatty acids or highly unsaturated fatty acids lower serum cholesterol in man For reasons indicated above I am more interested in the fundamental process or processes concerned with deposition of cholesterol and phospholipid in avascular tissues than with the level of serum cholesterol the rat deficient in EFA accumulates cholesterol and phospholipid in its epidermis although the serum cholesterol is abnormally low Many persons develop coronary thrombosis with normal serum cholesterol in the series published by James *et al* (1957) there was no difference between those who had had a myocardial infarction and the controls Conversely many persons are walking the streets with serum cholersterol levels fantastically elevated The work of Kinsell and others lends strong support to the thesis that EFA or polyethenoid fatty acids are required for the normal metabolism

in children may be enhanced if cod liver oil is also administered It is interesting that the disease is very rare in the U S A where either evaporated milks (in which EFA are more stable) or filled dried milks are often used the latter being skim milk to which vegetable seed oil—rich in linoleic acid—is added

Infants with idiopathic hypercalcaemia often show patchy fragmentation of the internal elastic lamina of smaller arteries with occasional impregnation with calcium salts As will now be discussed deficiency of EFA causes faulty connective tissue and this pathological change might be a direct result of such deficiency It is tempting to suppose that this might be relevant to the apparent increasing incidence of dissecting aneurysm of the aorta and that Monckeberg s sclerosis might be caused by deficiency of EFA with a relatively high intake of calcium and vitamin D

STRUCTURAL DEFECTS AS A BASIS OF DISEASE

Lower animals chronically deficient in EFA may appear on superficial examination to be fairly normal yet they react violently to insults or stresses For instance application to the back of one paw of a mild dose of ultra violet light that would produce no obvious lesion in a control rat may cause so violent a reaction in an EFA deficient rat that there is necrosis of much of the limb Ultra violet light destroys EFA and the permeability of the epidermis to water already excessive in deficiency of EFA is further increased by irradiation with UVL The dermatitis following UVL may be caused by destruction of EFA in the superficial layers of the epidermis with production of a product such as a hydroperoxide that causes a chain reaction of destruction of EFA and produces the erythema and necrosis The relation of EFA to detergent dermatitis and nickel dermatitis have been discussed elsewhere (Sinclair 1957b) The violent reactions to insults are similar to the responses found in the so called collagen diseases of man, or diseases of mesenchymal ground substance The lower animal deficient in EFA or pyridoxin may get duodenal ulcers ulcerative colitis nephrosis with haematuria easy rupture of capillaries arthritis osteoporosis and dental caries Both duodenal ulcers now a disease

ordinary linoleic acid as compared with human milk, but it is rich in calcium When it is dried the small amount of EFA in it is unstable in presence of air probably because metal ions, un hydrated, catalyse the oxidation of polyethenoid acids with formation first of hydroperoxides and later of aldehydes which cause it to smell and taste rancid An infant fed on National Dried Milk may therefore be on a diet rich in calcium and vita min D but low in EFA The deficiency of EFA could cause the rise in serum cholesterol (mainly esterified) and the alleged atheroma on lines already discussed Vitamin D has an esterifiable 3 hydroxyl group chemically very similar to that in choles terol By analogy, vitamin D might therefore normally be esterified in the body with EFA and its concentration, like that of esterified cholesterol, might be increased if it were esterified with unusual fatty acids in presence of a relative deficiency of EFA, but it might still exert its calcium retaining effect in the body In this way deficiency of EFA could cause hypersensitivity to vitamin D corn oil, a rich source of ordinary linoleic acid, should decrease the potency of vitamin D whereas cow s milk, a very poor source of ordinary linoleic acid, should increase the potency, and this is in fact the case (Lewis, 1935) The failure to grow, the polydipsia and the mental retardation might all be directly attributable to deficiency of EFA Infection increases the requirement of EFA and so would tend to precipitate defi ciency An attempt might be made as follows to explain the increase in plasma vitamin A This exists in plasma mainly as alcohol and is stored in liver mainly as ester, but no informa tion seems available about the fatty acids with which vitamin A is normally esterified If these were EFA it is possible that in deficiency storage would be hindered and the concentration of the alcohol in plasma increased To test this hypothesis of in creased sensitivity to vitamin D in deficiency of EFA, I ad ministered a non toxic daily dose of the vitamin to EFA deficient and control rats The former rapidly lost weight and died with uraemia caused by nephrocalcinosis whereas the con trol animals were unaffected even when given two times that dose of vitamin D There is therefore increased sensitivity to vitamin D in presence of deficiency of EFA and this deficiency

obtained upon these may be irrelevant to clinical science But the observational method and the epidemiological approach can on occasion help There are many diseases mainly chronic degenerative diseases, that are found predominantly in the more highly civilized countries, such as coronary thrombosis duo denal ulcers, dental caries, bronchial asthma and senile osteo porosis Many such diseases are becoming rapidly more preva lent in these countries and this is not explicable in terms of the increasing expectation of life indeed the expectation of life of a middle aged man in this country is only four years greater now than it was a century ago despite the tremendous achieve ments of medical science and this is because the degenerative diseases are increasing so fast It is tempting to seek a unified hypothesis for this phenomenon and twenty years ago I became interested in deficiency of essential fatty acids as a possible cause The evidence fits together, though the gaps in it are obvious and very wide One would prefer to try to fill those gaps before speculating unduly, for the scientific basis of medicine must rest upon verifiable facts but facilities are needed that I do not possess and I therefore commend to others the study of deficiency of essential fatty acids in man as a problem of great and accelerating urgency as our foods become increasingly processed

REFERENCES

AHRENS E H Jun INSULL W Jun BLOMSTRAND R HIRSCH J TSALTAS T T and PETERSON M L (1957) *Lancet* 1 943

BASNAYAKE V and SINCLAIR H M (1954) *J Physiol* 126 55P

BASNAYAKE V and SINCLAIR H M (1955) *Proc II Int Conf Biochem Probl Lipids* (Ghent) 4,6–484

BONHAM CARTER R E DENT C E FOWLER D I and HARPER C M (1955) *Arch Dis Childh* 30 399

BRONTE STEWART B (1957) *Nutr* 11 60

EVANS H M and BURR G O (1928) *Proc Soc exp Biol N Y* 25 390

HANSEN A E ADAM D J D BOELSCHE A N HAGGARD M E and WIESE H F (1957) *Proc IV Int Conf Biochem Probl Lipids* (Oxford) In press

HARTROFT W S and THOMAS W A (1957) *J Amer med Ass* 164 1899

of middle aged men, and senile osteoporosis are increasing rapidly in this country and dental caries is notoriously a disease of the more highly civilized countries where processed foods are eaten The altered structure of epithelium in the EFA deficient animal allows substances to penetrate more easily into the body, and infections are common This may be relevant to allergic conditions such as asthma in man in which foreign proteins penetrate the mucosa, the more easy passage of viruses through the nasal or intestinal mucosa might be of significance in the occurrence of coryza and the increasing prevalence of polio myelitis in more highly civilized countries Defective mito-chondrial structure might explain the preliminary evidence I have obtained that rats deficient in EFA are more prone to develop carcinoma of the stomach when fed methylcholanthrene than are controls If there is an increased sensitivity to ad ministered chemical carcinogens in EFA deficiency in man, one might expect carcinoma of stomach, bladder and lung (assum ing inhalation of a carcinogen) to be more common in males since their requirement of EFA is greater than that of females This possibility of deficiency of EFA playing a part in the cause of bronchial carcinoma might explain the prevalence of athe roma in this condition and also the nervous system changes which may also accompany carcinoma of the stomach The nervous system tends to retain EFA even when other tissues have become depleted but there is a tendency for depletion to occur As Macmillan and I have recently reported (1957), implanta tion of a tumour in an EFA deficient animal precipitates severe deficiency The cause of carcinomatous neuropathy might there fore be as follows EFA deficiency tends to increase sensitivity to a chemical carcinogen and to deplete the nervous system of EFA, if a carcinoma is produced it will quickly increase the severity of the deficiency and so precipitate lesions in the ner vous system perhaps in peripheral nerves, perhaps in the cere bellum or perhaps in the cortex

Conclusion

The scientific basis of medicine rests upon the application of the experimental method to man and lower animals, though results

obtained upon these may be irrelevant to clinical science But the observational method and the epidemiological approach can on occasion help There are many diseases, mainly chronic degenerative diseases, that are found predominantly in the more highly civilized countries such as coronary thrombosis, duodenal ulcers, dental caries bronchial asthma and senile osteoporosis Many such diseases are becoming rapidly more prevalent in these countries and this is not explicable in terms of the increasing expectation of life indeed the expectation of life of a middle aged man in this country is only four years greater now than it was a century ago despite the tremendous achievements of medical science and this is because the degenerative diseases are increasing so fast It is tempting to seek a unified hypothesis for this phenomenon, and twenty years ago I became interested in deficiency of essential fatty acids as a possible cause The evidence fits together, though the gaps in it are obvious and very wide One would prefer to try to fill those gaps before speculating unduly, for the scientific basis of medicine must rest upon verifiable facts but facilities are needed that I do not possess and I therefore commend to others the study of deficiency of essential fatty acids in man as a problem of great and accelerating urgency as our foods become increasingly processed

REFERENCES

AHRENS E H Jun INSULL W Jun BLOMSTRAND R HIRSCH J TSALTAS T T and PETERSON M L (1957) *Lancet* **1** 943

BASNAYAKE V and SINCLAIR H M (1954) *J Physiol* **126** 55P

BASNAYAKE V and SINCLAIR H M (1955) *Proc II Int Conf Biochem Probl Lipids* (Ghent) 476–484

BONHAM CARTER R E DENT C E FOWLER D I and HARPER C M (1955) *Arch Dis Childh* **30** 399

BRONTE STEWART B (1957) *Nutr* **11** 60

EVANS H M and BURR G O (1928) *Proc Soc exp Biol N Y* **25** 390

HANSEN A E ADAM D J D BOELSCHE A N HAGGARD M E and WIESE H F (1957) *Proc IV Int Conf Biochem Probl Lipids* (Oxford) In press

HARTROFT W S and THOMAS W A (1957) *J Amer med Ass* **164** 1899

HOLLINGSWORTH, D F VAUGHAN M C and WARNOCK, G M (1956) *Proc Nutr Soc* 15 xvii
JAMES A T LOVELOCK, J E WEBB J and TROTTER W R (1957) *Lancet* i 705
KEYS A (1953) *J Mt Sinai Hosp* 20 134
KINSELL, L W, MICHAELS G D and DAILEY, J (1957) *Proc IV Int Conf Biochem Probl Lipids* (Oxford) In press
KINSELL L W and SINCLAIR, H M (1957) *Lancet* i 883
KLEIN, P D and JOHNSON R M (1954) *J biol Chem* 211 103
KNISELY M H (1951) *Postgrad Med* 10 15
KRAMÁR, J and LEVINE, V E (1953) *J Nutr* 50 149
LEWIS, J M (1935) *J Pediatr* 6 362
LIGHTWOOD R (1952) *Proc roy Soc Med* 45 401
MABROUK, A F and BROWN, J B (1956) *J Amer oil Chem Soc* 33 98
MCCANN M B and TRULSON M F (1957) *J Amer diet Ass* 33 358
MACMILLAN A L and SINCLAIR, H M (1957) *Proc IV Int Conf Biochem Probl Lipids* (Oxford) In press
MALMROS H and WIGAND, G (1957) *Lancet* ii, 1
MEAD J F and HOWTON D R (1957) *Proc IV Int Conf Biochem Probl Lipids* (Oxford) In press
PASSMORE R (1953) *Lancet* i 639
RAMALINGASWAMI V and SINCLAIR H M (1953) *Brit J Derm* 65 1
ROBINSON D S and POOLE J C F (1956) *Quart J exp Physiol* 41 36
SINCLAIR H M (1952) *Biochem Soc Symp* No 9 80
SINCLAIR H M (1956a) *Lancet* i 381
SINCLAIR H M (1956b) *Lancet* ii 101
SINCLAIR H M (1956c) *Lancet* ii 693
SINCLAIR H M (1957a) *Proc III Int Conf Biochem Probl Lipids* (Brussels) 392–400
SINCLAIR H M (1957b) *Ann Nutr* (Paris) 11 A147
SINCLAIR H M (1957c) *Roy Soc Hlth J* 77 234
SINCLAIR H M (1957d) *Proc Nutr Soc* In press
SINCLAIR H M (1957e) *Proc IV Jorn Diet Madrid* 179
SREENIVASAN B and BROWN J B (1956) *J Amer oil Chem Soc* 33 521
WITTEN P W and HOLMAN R T (1952) *Arch Biochem* 41 266
YERUSHALMY J and HILLEBOE H E (1957) *N Y State J Med* 57 2343
YUDKIN J (1957) *Lancet* ii 155

XVI

The Control of Fat Metabolism

A L GREENBAUM

The physical appearance of patients suffering from thyrogenic or androgenic adiposity or resulting from Frohlich's or Cushing s disease emphasizes the profound effect that hormones can exert on the course of fat metabolism and of course, the hormones are not the only means which produce such a derangement In addition to obvious factors like diet and nutritional status there are heredity hypothalamic lesions and other clinical conditions In this lecture however the clinical aspects of fat metabolism will not be considered at all but rather the possible regulatory mechanisms of fat oxida tion and synthesis will be sketched out from a biochemical point of view in the hope that clinicians may find in the imbal ance of one or more of these mechanisms the biogenesis of the pathological condition

Figure 1 gives a very brief outline of the interrelationship of fat and carbohydrate metabolisms

The common meeting point of the fat and carbohydrate metabolic pathways is at the level of the two-carbon fragment acetyl coenzyme A This unit arises from the oxidation of both fats and carbohydrates and it is with the metabolic fate of this fragment that this lecture is largely concerned Under normal circumstances most of it will be oxidized to carbon dioxide and water by way of the Krebs tricarboxylic acid cycle (TCA) and in the process a great deal of energy will be liberated Some of it will be synthesized to fatty acids irrespective of whether its original source was a carbohydrate or another fatty acid Very little, if any, will form acetoacetate Acetoacetate production in

quantity will only occur when the subsequent metabolism of the acetyl coenzyme A in the Krebs cycle is blocked in any way and may be regarded as an entirely abnormal product

For the purpose of the present discussion it will be taken that the production and the subsequent metabolic behaviour of the

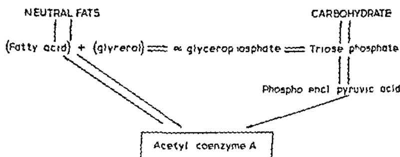


Fig 1

acetyl coenzyme A is regulated by enzymic and hormonal mechanisms and we shall consider the factors which affect its metabolism

The mechanisms by which acetyl coenzyme A is produced from carbohydrate are too well known to justify a detailed account here, but its production from fatty acids has only recently been fully worked out In this connection it does not matter whether the fatty acids are derived from neutral fat or from phospholipids both of which can yield them, the scheme is a general one

As a result of the work of Green and Lynen and their schools in the U S A and Germany the mechanism of fatty acid oxidation and synthesis is known in some detail and this, in turn, permits us to visualize some of the factors which can influence not only the rate of fatty acid metabolism but also the direction, that is to say, whether the emphasis will be on fatty acid oxidation or on its reverse, fatty acid synthesis The full details of fatty acid oxidation are given in Figure 2

After a preliminary activation the fatty acids are oxidized by the repetition of a cycle of four reactions The exact sequence of reactions in the activation procedure is not fully

known There are at least three enzymes catalysing this step which activate fatty acids of different chain lengths According to Berg (1955) the activation of acetate involves first, the inter action of the acetate with adenosine triphosphate (ATP) with the formation of an acetate nucleotide complex and then the

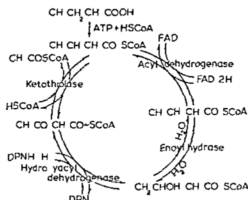


Fig 2 The fatty acid cycle ATP=adenosine triphosphate HSCoA =coenzyme A FAD=flavin adenine-dinucleotide DPN=diphosphopyridine nucleotide

replacement of the nucleotide component by coenzyme A so that acetyl coenzyme A is formed It is supposed that a similar sequence occurs in the activation of higher fatty acids which consequently form acyl coenzyme A derivatives It is these derivatives which form the substrate for the fatty acid cycle

The first reaction of the cycle is a dehydrogenation removing two hydrogens and giving rise to an unsaturated acyl coenzyme A Secondly there is a hydration of this unsaturated compound by the addition of the elements of water across the double bond to produce a hydroxy acyl coenzyme A and then the third reaction which follows is another dehydrogenation The removal of this second pair of hydrogens leaves a β keto acid The final reaction of the cycle is the cleavage of the keto acid at the β carbon by another molecule of coenzyme A so that the last two carbons of the chain are split off as a two-carbon fragment and the residual chain is left activated ready for a further turn

round the cycle The two carbon unit which is split off is acetyl coenzyme A By a repetition of this cycle of reactions the fatty acid chain is eroded away, two carbons at a time, until the whole fatty acid is degraded to a series of acetyl coenzyme A units

The scheme shown in Figure 2 gives the details of this re action sequence complete with all the enzymes and coenzymes and illustrates several points which are worth emphasizing Firstly, it should be noted that the cycle is fully reversible the same four enzymes can either degrade or synthesize fatty acids During degradation the cycle is running in an oxidative manner, that is to say, it is removing hydrogens from the substrate, while during synthesis it is operating as a reductive cycle, that is, adding hydrogens to the substrate to form the hydrogen rich fatty acids The second point is that there are two steps in this cycle which require the presence of coenzyme A, the initial activation and the final step by which the acetyl coenzyme A is cleaved off and the residual chain reactivated ready for another turn Thirdly, there are two steps involving the enzymic transfer of hydrogens, the first is the acyl dehydrogenase which, in the oxidative direction takes two hydrogens from the saturated fatty acid and transfers them to flavin adenine dinucleotide (FAD) The second is the β hydroxyacyl dehydrogenase which removes hydrogens from the hydroxy acyl compound and passes them to the diphosphopyridine nucleotide (DPN) In the reverse direction, that is during fatty acid synthesis the same two en zymes act reductively, the hydrogens being derived from the reduced forms of the coenzymes, FADH and DPNH re spectively There are two final points to be made which are not emphasized in Figure 2 The first of these is that for syn thesis to occur we must have an input of energy All synthetic processes in the body are energy requiring reactions and fatty acid synthesis is no exception The second point is also related to energy relationships

Energy Acceptors and Donators

In biochemical reactions involving a change of energy level there is normally present an energy acceptor if the reaction is exergonic or an energy donator if the reaction is endergonic

The acceptor is usually adenosine diphosphate (ADP) and the energy donator is adenosine triphosphate (ATP) The energy exchange is accompanied by the transformation of one of these compounds into the other



ADP+inorganic ℗ + e ⇌ATP

Energy yielding reactions *require* the presence of ADP to act as an acceptor and in the case of some of the reactions we are discussing here ADP is an obligatory component of the reaction In reactions of this kind the amount of ADP present can be a limiting component of the reaction mixture and in such circumstances the amount of ADP available will be the rate determining factor Similar considerations apply to other co factors like DPN which can also act as a rate determining factor Rate regulation by co-factor levels profoundly influences the metabolism of the fatty acids and it would be profitable to consider in a more detailed way the effect of these various co-factors on the course of the reactions which were outlined earlier

The first of these co factors is coenzyme A It is possible to induce a coenzyme A deficiency in an animal by maintaining it on a diet deficient in pantothenic acid which is a precursor of coenzyme A In such an animal fat synthesis is practically absent If now pantothenic acid is fed to such a deficient animal then as Klein and Lipmann (1953) have shown there is a strict parallelism between the level of coenzyme A appearing in the liver and the rate at which the liver can synthesize lipid Or again Lotspeich (1950) has induced a pantothenic acid deficiency in some hormone treated animals and shown that the amount of coenzyme A in these animals could be rate limiting for fat metabolism But it must be remembered that it is almost impossible to obtain a pantothenic acid deficiency in an animal on a reasonably normal diet We might expect then that normal animals would always have adequate coenzyme A precursor and that there was little likelihood of a deficiency occurring to act as a regulatory factor But while we can discard the idea that a coenzyme A deficiency is a regulator I must point out that there is real possibility that high levels of

this coenzyme can act in this capacity Brady, Mamoon & Stadtman (1956) have shown in rats that high levels of liver coenzyme A favour fat oxidation and inhibit fat synthesis and, in confirmation of this, Green (1956) has reported that the optimal concentration of coenzyme A for fat synthesis is very low, higher concentrations being completely inhibitory At the moment both mechanism and the physiological significance of this finding are obscure, but the fact remains that if these results do apply to the whole animal then some control can be exerted on the rate of fat oxidation and synthesis by the level of coenzyme A in the liver It is interesting to notice in this connection that Weiland, Reinwein and Lynen (1956) found 35 per cent more coenzyme A in the livers of diabetic rats than in their controls which, on this formulation, would have the effect of promoting fat oxidation and inhibiting fat synthesis, and this corresponds very well with what actually occurs in a diabetic liver

The second co factor we must discuss is adenosine diphosphate and the level of this can most certainly play a regulatory role much as it does in carbohydrate metabolism As I mentioned earlier, if ADP is an obligatory acceptor for an exergonic reaction then the amount of available ADP will be a rate limiting factor There are two steps in fat oxidation, the two dehydrogenations in which the subsequent oxidation of the hydrogens removed is coupled to a phosphorylation requiring the presence of ADP When the level of ADP is low, and this would be the case in a body at rest then the rate of hydrogen oxidation is low and so, moving further back along the sequence, is the oxidation of fatty acids But as soon as any physiological activity occurs there is a breakdown of some ATP to ADP to provide the energy for this activity The ADP so formed can now act as a phosphate acceptor for the coupled phosphorylation which is associated with hydrogen oxidation and the rate of this reaction is consequently increased and so, also, is fat oxidation This stimulating effect of ADP on respiration has been amply demonstrated by Britton Chance in his steady state measurements (see Chance, 1953-4) We have here an extremely efficient control mechanism where the rate of fat oxidation is directly geared to the amount of work done A similar mechanism

operates of course for carbohydrate metabolism which also has such ADP linked reactions, and this brings us to a second controlling mechanism associated with ADP

The amount of ADP in the tissues is not unlimited But this same amount, small as it is, is necessary for both fat and carbohydrate metabolism Unless the two metabolisms have their own individual supply of ADP it follows that there will be a competition for the available ADP by fat metabolism on the one hand and carbohydrate metabolism on the other The faster carbohydrate oxidation goes the less ADP there will be available for fat oxidation and vice versa The two systems would be, therefore, not exactly mutually exclusive but at least inversely related That such a competition does actually exist *in vivo* is shown by the studies of Lossow and Chaikoff (1955) in which they have shown that rats forcibly fed with carbohydrate have a decreased capacity for fat oxidation

Direction of the Fatty Acid Cycle

The foregoing paragraphs have been a discussion of some of the factors which control the *rate* of fat oxidation but we must now turn to another aspect of fatty acid metabolism According to Figure 2 the fatty acid cycle is completely reversible the same enzymes catalysing either synthesis or oxidation Obviously some form of control is necessary here to impose a direction of turning on the cycle It would be chaotic if some of the enzymes were operating in a synthetic direction and others oxidatively controlled only by the reaction equilibria

There are three main points where the direction of the cycle can be determined The first of these is at the β hydroxy acyl dehydrogenase stage where the availability of DPN or DPNH will determine whether the cycle operates in an oxidative or reductive direction The second point is at the acyl dehydro genase step where presumably the availability of oxidized or reduced flavin adenine dinucleotide could also act as a direction-determining factor in much the same way as DPN Thirdly a kinetic factor has been postulated operating to increase the rate of synthesis by the removal of the end products of synthesis If these end products accumulate that is the newly formed fatty

acids, then the whole reaction sequence will be slowed according to the laws of mass action The reaction affected by this kinetic argument is the fourth step in the cycle, that brought about by the enzyme keto acyl thiolase At this step another possible controlling factor would be the presence of glycerol which is needed for the removal of the completed fatty acids by esterification

We will now consider these points in some detail

The first point was the effect of the DPN DPNH ratio on the direction of the β hydroxy acyl dehydrogenase This particular reaction is the most thoroughly investigated and best understood of the whole fatty acid cycle sequence I will not quote the many studies of this reaction but I must refer to the work of Hele, Popjak and Lauryssens (1957) The studies of these authors stem from the findings of Popjak and Tietz (1955) that a soluble fraction from the homogenized mammary gland contains the whole complex of fatty acid synthesizing enzymes. The synthesis of fatty acid from acetate precursor is illustrated in Figure 3

It can be seen from this figure that the synthesis of fatty acid follows hand in hand the oxidation of DPNH When the synthesis of fatty acid stops it can be restarted, explosively as the authors put it by the addition of fresh DPNH The interesting point about this figure is that fatty acid synthesis stops when only about one half of the DPNH present has been oxidized This figure of only about one half oxidized is due to the setting up of an equilibrium state for the reaction catalysed by the enzyme β hydroxy acyl dehydrogenase We may say, therefore that we need a considerable pressure of DPNH to drive the fatty acid cycle in a synthetic direction The work of Hele *et al* (1957) supports an idea put forward by Lynen (1953) some four years ago in which he said It would appear that in the living cell the DPN DPNH ratio may determine whether synthesis or degradation of the carbon chain may occur' If this view is correct and the ratio of oxidized to reduced nucleotide is a rate and direction determinant then we should consider the factors which influence the production and oxidation of DPNH We cannot of course, consider this in any detail in the

confines of this short lecture and it must suffice if I produce the simplification that the main sources of DPNH apart from the fatty acid cycle itself are, in the main the enzymes of the glycolytic pathway and the various dehydrogenases associated with the Krebs tricarboxylic acid cycle There is reason to

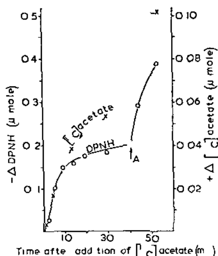


FIG 3 The utilization of DPNH and the incorporation of [14c] acetate into fatty acids A represents the point at which further DPNH was added (Reprinted with permission from Hele Popják and Lauryssens 1957)

believe that the DPNH produced from these two sources is not of equal value for fatty acid synthesis and that the main source of DPNH for this purpose is from the glycolytic enzymes In this connection I shall only quote one paper that of Br ind and Helmreich (1956), in which they show a close parallelism between the amount of fatty acid synthesized and the activity of the glycolytic system under a variety of conditions They could find no relationship between fatty acid synthesis and respiration and respiration may be taken as a measure of the activity of the Krebs cycle

The importance of DPNH generated by the enzymes of the glycolytic sequence as a driving force for fatty acid synthesis is emphasized even more strongly if we consider the effect of

starvation or of the hormone insulin on fat synthesis The effect of starvation on lipogenesis has been studied by Chaikoff and his associates in a long series of papers In summary their results show that while the liver of a starving rat retains its power to oxidize fats at a normal rate, it has almost lost its capacity to synthesize fatty acids from acetate precursor Injection of in sulin could not restore the power of lipogenesis from acetate, but the feeding of glucose could When insulin and glucose were given together lipogenesis was raised to supernormal levels As starvation was found to have a profound effect on carbohydrate metabolism which could not be reversed by insulin but could be by glucose, they concluded that an active glycolysis was necessary for the incorporation of acetate into fatty acids The effect they had observed after feeding glucose, or giving glucose plus insulin was thought to be simply due to the priming of carbohydrate metabolism

Even more striking results were obtained by the Chaikoff school using the livers of diabetic rats The liver of the diabetic rat has lost nearly all capacity to synthesize fat from either glu cose or acetate although here again the ability to oxidize fats is entirely unimpaired Lipogenesis from either of these sub stances was restored by insulin treatment and, in addition, fructose feeding permitted the utilization of acetate for fat synthesis

These results were interpreted as indicating two blocks in lipogenesis The first of these was at the level of the enzyme hexokinase and this could be cured only by the administration of insulin The second block was held to occur beyond the hexo-kinase reaction, probably at the level of the two carbon frag ment This second block could be relieved by fructose feeding as well as by insulin While the mechanism of the action of fructose is not yet fully explained it is generally believed that its virtue lies in the fact that its use permits glycolytic enzymes to produce DPNH and it is this DPNH which permits fat syn thesis A similar line of reasoning leads to the view that the supernormal rate of lipogenesis found in starving rats given insulin plus glucose is due to the increased rate of DPNH forma tion consequent upon the increased activity of the glycolytic

enzymes We have, therefore, good evidence that the DPN DPNH ratio affects both the rate and the direction of the enzyme hydroxy acyl dehydrogenase The second point we mentioned at which control of fatty acid synthesis could be effected was at the acyl dehydrogenase step where the availability of oxidized and reduced flavin adenine dinucleotide would be the factor involved I should mention here that the flavin adenine di nucleotide is built into the enzyme itself and that three flavo protein acyl dehydrogenases, each working on fatty acids of different chain lengths, have been isolated The reaction in which these flavoprotein enzymes participate is either the removal or addition of hydrogens in the first reaction of the cycle Very little is known about this particular reaction although it is supposed that the same considerations apply here as applied to the DPN linked step we have already dis cussed The situation is made somewhat more complicated by the nature of the electron transferring factor described by Beinert and his associates (Crane and Beinert, 1956) This factor is supposed to cause oxidation by removing hydrogens from the reduced flavin adenine dinucleotide of the acyl de hydrogenases The reverse of this reaction, that is the reduction of the oxidized flavin adenine dinucleotide which would of course be a necessary preliminary for fatty acid synthesis has not yet been demonstrated But although it would seem logical to believe that this reversal is possible the immediate source of these hydrogens for the reduction is still in doubt It has been generally assumed that the source is again DPNH and if this were so then at this point also the DPN DPNH ratio would act as a controlling factor in the way we have just considered But this is by no means certain and I shall return to this point later

Our third method of control was the possibility of some form of kinetic control at the last reaction of the cycle Each turn of the cycle elongates the fatty acid chain by two carbons Clearly there must be some limit to the process or we should end up with fatty acids of infinite chain length The factor which arbitrates whether a ketoacyl coenzyme A compound should remain within the cycle and be further elongated or whether it should

be passed out of the cycle as a finished fatty acid, is the affinity of the enzyme β ketoacyl thiolase for its substrate, the ketoacyl coenzyme A in question If the affinity of this enzyme is high, further elongation will occur, if it is low then the fatty acid will be liberated Although it now seems probable that there are two or three thiolases, the thiolase isolated by Goldman (1954) has a high affinity for ketoacyl coenzyme A compounds of short chain lengths but this declines precipitously as the chain gets *to 14* to 16 carbons and *is virtually zero for chain lengths of* 18 carbons At this point, therefore, the fatty acid becomes detached from the enzyme and elongation of the chain ceases

The normal fate of these detached acyl coenzyme A derivatives is for them to become esterified with α glycerophosphate to form either neutral fats or phospholipins The enzyme catalysing this step is the one discovered by Kornberg and Pricer (1953) It has been suggested that the activity of this enzyme could be a regulatory factor because it removes the end product of the action of the enzyme ketothiolase and so prevents an accumulation of this end product to such an extent that the ketothiolase reaction, and through it the whole fatty acid cycle, would be slowed by mass action behaviour This seems to me unlikely Since ketothiolase has such a low affinity for the acyl coenzyme A compounds which are involved in the Kornberg and Pricer reaction it would need extremely large changes in the activity of this esterifying enzyme before they reflected in changes *in the rate of fatty acid synthesis So far there* is no quantitative evidence to show that any drastic changes in the activity of this enzyme occur But even without any change in enzyme activity some kinetic control can be envisaged at this same reaction if the other component of the esterifying system, that is glycerol, were deficient and there does appear to be some evidence that glycerol can indeed be a rate limiting component in fat synthesis It is not easy to assess how important a factor glycerol lack might be *in vivo* when the glycolytic enzyme system, which was earlier described as being necessary for fatty acid synthesis, would also automatically provide α glycerophosphate for the Kornberg and Pricer enzyme But although this would be true for normal animals it is almost certainly not so for

diabetic ones Nor would it be true of those animals in which glucose is not the main metabolic fuel the ruminants for instance In these cases, as French and Popjak (1951) have shown glycerol lack may well be a limiting factor in fat synthesis Nevertheless all that we can say at the moment is that while the system removing fatty acids by esterification could possibly act in controlling the rate of fatty acid synthesis, it is probably not so important as the other points which have been mentioned

THE RATIO OF OXIDIZED TO REDUCED DPN

I want to return now to a point I mentioned earlier the influence of the DPN DPNH ratio on the rate and direction of fatty acid synthesis One of the difficulties in visualizing how a tissue changes over from oxidizing fatty acids to synthesizing them is in the expected change there would have to be in the DPN DPNH ratio Such measurements as have been made of the ratio in the whole cell more or less agree that there is about three times as much DPN as DPNH and yet on the basis of the work by Hele Popják and Lauryssens (1957), fat synthesis requires that there should be more DPNH than DPN Changes of the DPN DPNH ratio of the kind that their work shows to be required for fat synthesis have never been found experimentally The possibility exists that as the system on which Hele *et al* (1957) work is present in the supernatant fraction it would be sufficient if a high DPNH to DPN ratio existed in the supernatant but not necessarily in the cell as a whole But here again there is something of a quandary because the only measurements of the DPN DPNH ratio in the supernatant are those of Glock and McLean (1956) and these show that while the ratio for the whole cell is three to one in favour of DPN in the soluble fraction it is seven to one, conditions entirely against fatty acid synthesis Several possible explanations can be put forward to account for the fact that fat synthesis can go on in spite of this adverse ratio Firstly, one can postulate that fat synthesis does not occur in the whole cytoplasm but only in a small part of it a localized area like the Golgi region for instance Electron microscopy has shown that the cytoplasm which was once regarded as a sort of homogeneous matrix with

the enzymes swimming about in it, in reality has a most complicated structure in which it would be quite conceivable to have discrete zones It would then be possible to have a limited region with a ratio of oxidized to reduced nucleotide favouring fat synthesis while the cell as a whole had a ratio against it While such an idea may appear possible there is, so far as I am aware, no real evidence for it and it must be left simply as a possibility

Secondly, it could be that DPNH is not the only factor concerned The work which really implicates diphosphopyridine nucleotide is mostly *in vitro* work with purified enzymes *In vivo* other factors may be involved As I shall show later quite a case can be made for another regulatory factor acting in much the same way as DPNH Thirdly, there is another aspect of this co factor story which we have not considered yet In the last few years it has become increasingly obvious that many, if not most, of the enzymes involved in hydrogen transfer have their coenzymes very firmly attached to them, their own private coenzyme so to speak, which need not be in the same state of oxidation or reduction as another molecule of the same coenzyme attached to a different enzyme It would then be possible to have the hydroxy acyl dehydrogenase bound DPN mainly in the reduced form while the rest of the cellular DPN is mainly in the oxidized form Again, if this were so fat synthesis would be possible in conditions where normal analytical methods would show it to be unlikely In fact, Holzer, Schultz and Lynen (1956), have shown that very little DPNH exists free in the cytoplasm and in yeast it is probably of the order of 1 μmole DPNH *per litre* of cytoplasm

This idea of DPNH localization would go far to explain the position that has arisen in studies of the total cellular DPN and DPNH in the livers of hormonally treated rats To illustrate how difficult is the interpretation of such overall figures I will quote some experiments done in my own laboratory We have studied this ratio in two conditions the first in insulin treated rats where Chaikoff has shown a considerable acceleration of fat synthesis (see Chaikoff 1951–2) and the second in rats treated with pituitary growth hormone This second hormone

has a profound depressent effect on the rate of fatty acid synthesis (see Table 1)

The figures quoted are taken from some work done in collaboration with Dr Glascock (1957) The substrates used were [carbonyl ^{14}C] pyruvate and [carboxy ^{14}C] acetate The

TABLE 1 *The rate of incorporation of labelled substrates into the fatty acids* and phospholipins of surviving liver slices from growth hormone treated rats *

	Fatty acids (Counts/min /mg C)	Phospholipids (Counts/min /mg C)
A [*carbonyl*] ^{14}C-pyruvate	41 0±10 1	73 0±7 5
B [*carboxy*] ^{14}C acetate	39 3±12 3	54 2±7 2

The results are expressed as a percentage of the incorporation rate into the same lipids of liver slices from an untreated rat run at the same time

* Animals injected with 1 mg purified growth hormone six hours before killing

figures are expressed as percentages of the control values The incorporation rate of the labelled carbon into the fatty acid of the treated rats fell to as little as 40 per cent of the control rate and phospholipin synthesis was halved At longer time intervals (up to 12 hours) the values were as low as 20 per cent Thus, the two hormonal conditions we have used produce diametrically opposed effects on fat metabolism an intense stimulation of fat synthesis by insulin and an intense inhibition of fat synthesis by growth hormone We expected that such different metabolic patterns would be reflected in interesting differences in the cellular DPN and DPNH a low ratio in insulin treatment and a high ratio in growth hormone treatment Table 2 shows the results of this experiment (Greenbaum and Graymore 1956)

As was expected insulin markedly lowered the DPN DPNH ratio and it did so by greatly increasing the amount of DPNH rather than by reducing DPN The ratio appeared to bear out Lynen s suggestion that a high DPNH was needed for fat synthesis The growth hormone results were however most unexpected for here too there is a reduced DPN DPNH ratio and here also this is mainly due to an increase in the DPNH Although the table shows the result of growth hormone treatment for only one time interval it should be added here that

very similar results were obtained at all the time intervals tested, from as little as one hour to as long as seven days We have here a shift in DPN DPNH ratio to one favourable to fat synthesis in a rat which not only has a reduced rate of fat synthesis (see Table 1), but which is undoubtedly oxidizing fat at

TABLE 2 Changes in the level of oxidized and reduced DPN in the livers of rats treated with insulin and pituitary growth hormone

Treatment	DPN (μg/g liver)	DPNH (μg/g liver)	DPN DPNH
A *Insulin*			
Controls	349±7 8	162±3 7	2 20
0 25 units insulin 1 hour before killing	406±9 1	306±22 5	1 35
Diabetic rats	346±17 4	99 8±5 8	3 48
B *Growth Hormone*			
Controls	399±16 2	213±12 6	1 91
1 mg GH injected 24 hours before killing	401±8 5	256±4 6	1 56

an accelerated rate (Greenbaum and McLean 1953) If the insulin results fitted in with Lynen's ideas, then the growth hormone results certainly did not

Actually I think the explanation of this anomaly is reasonably straightforward With insulin treatment the changes in the reduced DPN can be accounted for largely in terms of the effect of insulin on the glycolytic enzymes, which are probably the major source of the DPNH, in that the injection of insulin increases glycolysis But the question now arises as to the source of the extra DPNH in growth hormone treated rats It could hardly be from glycolysis because Recant (1952) has shown that rate of glycolysis is not affected by growth hormone It seems much more likely that the increased DPNH which we found represented not so much a potential for fatty acid synthesis, but the very opposite, the result of an increased oxidation of fatty acids Each turn of the fatty acid cycle leads to the reduction of one molecule of DPN by the hydroxyacyl dehydrogenase

These results emphasize that we have been interpreting the

idea of the DPN DPNH ratio in much too general terms Probably what we should be considering is the ratio of the oxidized to reduced DPN bound to the enzyme itself, and, of course, there have been no measurements of this at all

Flavin linked Reaction

We have considered so far only one of the two hydrogenation steps and we must now turn to the other the flavin linked reaction Langdon (1955) has described the preparation of an enzyme system in the supernatant fluid left after high speed centrifugation of a liver homogenate which can bring about the synthesis of fatty acids In this system full activity can only be obtained if there is present, in addition to diphosphopyridine nucleotide some reduced triphosphopyridine nucleotide (TPNH) He has also shown that the TPNH is concerned with the flavin linked hydrogenation that is the reduction of the unsaturated acyl coenzyme A compound to the saturated compound and that, further DPNH cannot substitute for TPNH in this reaction In this case the TPNH level would be a rate and direction determining factor for the flavin linked enzyme, just as DPNH is for the hydroxy acyl dehydrogenase There is actually some supporting evidence for such an idea First there is the finding of Glock and McLean (1956) that the supernatant from a homogenate contains about four times as much TPNH as TPN a ratio of the right order for fat synthesis Second there is the work of Crane, Hauge and Beinert (1955) which shows that the electron transferring factor, that is, the enzyme which feeds the hydrogens needed for the reduction of the unsaturated compound on to the flavin which is the immediate hydrogen donor, has a quite strong TPNH dehydrogenase activity Third, there is the finding of Brady Mamoon and Stadtman (1956) that the addition of citrate to a system synthesizing fatty acids has a greater stimulating effect than the addition of any of the other intermediates of the Krebs cycle These authors suggest that the greater effectiveness of citrate can be ascribed to the fact that its oxidation by isocitric dehydrogenase produces TPNH whereas the oxidation of the other members of the cycle produces DPNH The superior quality of isocitric

acid has also been confirmed in Green's laboratory (Gibson *et al*, 1957) although the Wisconsin workers believe that iso citric has another, specific action as well as acting as a generator of TPNH The final point which bears on this subject is this As shown earlier, the oxidation of glucose stimulates fatty acid synthesis and this has been thought to be due to its acting as a source of DPNH But it should be pointed out here that the oxidation of glucose not only produces DPNH through the activity of the glycolytic enzymes, but also TPNH through the activity of the enzymes of the pentose phosphate shunt, the alternative pathway for glucose oxidation, and this could be just as important as DPNH

In direct contrast to this work of Langdon (1955) we have the work of Popjak and his collaborators In Popjak s work the enzyme system used is also a high speed supernatant, but this time from mammary gland With this preparation fatty acid synthesis can be readily achieved with only DPNH present, and, in fact, in their latest paper Hele, Popjak and Lauryssens (1957) have definitely excluded TPNH as a hydrogen source for fatty acid synthesis in their system

It is not easy to reconcile the results of these workers with those of Langdon and for the moment one must fall back on the usual explanation in such circumstances and say that the difference might be explicable in terms of different coenzyme specificities of the enzymes of liver and mammary gland or perhaps of even rabbit and rat

We have considered fatty acid synthesis in soluble enzyme systems and I now just want to say a few words about the mitochondrial enzymes Although it has been shown that each of the enzymes of the fatty acid cycle can be isolated as a soluble enzyme, and each of these enzymes is reversible, it has not yet been possible to reconstruct a fatty acid cycle from the soluble enzymes of mitochondria which can carry the synthesis of fatty acids further than the four carbon stage, although such a re constructed system can readily oxidize fatty acids with chain lengths up to 18 carbon atoms It is possible that this difference could mean that the mitochondria are primarily oxidative with fat synthesis occurring in the supernatant fraction This is

perhaps, not so extraordinary a suggestion as it appears at first sight Although the fatty acid cycle is usually written as being fully reversible there is no direct, unambiguous evidence that it really is although I think that the balance of evidence indicates that it is What seems more probable is that the mitochondrial system so far isolated is incomplete and that there are more co factors to come It seems reasonably certain that the tally of enzymes and coenzymes involved in fatty acid synthesis is not yet complete

We have reviewed in some detail the role of the coenzymes in regulating fatty acid metabolism and I must now finally mention the only work I know concerning changes in the activity of the enzymes of the fatty acid cycle This is the work of Weiland, Reinwein and Lynen (1956) on the levels of the enzymes in the livers of diabetic rats No measurements were made of enoyl hydrase, but of the other three, ketoacyl thiolase was slightly increased in the diabetic liver and the hydroxyacyl dehydrogenase by as much as 40 per cent The activity of acyl dehydrogenase was increased by over 100 per cent These changes are adaptive changes which reflect the increased rate of fat catabolism in the diabetic liver

Conclusion

So we arrive at the final conclusion that in all probability there are many checks and controls on fat metabolism There is the immediate and direct action of factors like coenzyme A adenosine diphosphate and the ratio of oxidized to reduced nucleotides but other, more long term factors can operate also Indirect hormonal control, for instance not only by insulin but also by the hormones of the anterior lobe of the pituitary and from the adrenal cortex, may play a role as also can adaptive changes in the activity or level of the fatty acid cycle itself The list of factors which might affect fat metabolism is a long one and by its very length it emphasizes how new and unexplored is the subject of metabolic regulation in biochemistry

REFERENCES

BERG P (1955) *J Amer chem Soc* **77** 3163
BRADY R O MAMOON A M and STADTMAN E R (1956) *J biol Chem* **222** 795
BRAND V and HELMREICH E (1956) *Biochem Z* **328** 146
CHAIKOFF I L (1951–2) *Harvey Lectures* p 99
CHANCE B (1953–4) *Harvey Lectures* p 145
CRANE F L and BEINERT H (1956) *J biol Chem* **218** 717
CRANE F L HAUGE J G and BEINERT H (1955) *Biochim biophys acta* **17** 292
FRENCH T H and POPJÁK G (1951) *Biochem J* **49** p III
GIBSON D M JACOB M I PORTER, J W TIETZ A and WAKIL S J (1957) *Biochim biophys acta* **23** 219
GLOCK G E and MCLEAN P (1956) *Exp Cell Res* **11** 234
GOLDMAN D S (1954) *J biol Chem* **208** 345
GREEN D E (1956) In *Biochemical Problems of Lipids* Butterworth s Scientific Publications (Lond) Ed Popjak G and Le Breton E p 233
GREENBAUM A L and GLASCOCK R F (1957) *Biochem J* **67** 360
GREENBAUM A L and GRAYMORE C N (1956) *Biochem J* **63** 163
GREENBAUM A L and MCLEAN, P (1953) *Biochem J* **54** 413
HELE P POPJAK G and LAURYSSENS M (1957) *Biochem J* **65** 348
HOLZER H SCHULTZ G and LYNEN F (1956) *Biochem Z* **328** 252
KLEIN H P and LIPMANN F (1953) *J biol Chem* **203** 101
KORNBERG A and PRICER W E (1953) *J biol Chem* **204** 345
LANGDON R G (1955) *J Amer chem Soc* **77** 5190
LOSSOW W J and CHAIKOFF I L (1955) *Arch Biochem Biophys* **57** 23
LOTSPEICH W D (1950) *Proc Soc exp Biol Med* **73** 85
LYNEN F (1953) *Harvey Lectures* p 210
POPJÁK G and TIETZ A (1955) *Biochem J* **60** 147
RECANT L (1952) *Fed Proc* **11** 272
WEILAND O, REINWEIN D and LYNEN F (1956) In *Biochemical Problems of Lipids* Butterworth s Scientific Publications (Lond) Ed Popjak G and Le Breton E p 155

XVII

The Rôle of the Endocrine System in Breast Cancer

GEOFFREY HADFIELD

THERE is an intimate biological relationship between the ovaries and mammary glands Both organs develop *pari passu* up to sexual maturity and become atrophic after the menopause Ovarian grafts restore the structure and function of the atrophic mammary glands which result from bilateral oophorectomy (Gregoriev 1897) and also induce a growth response in the rudimentary mammary glands of males (Steinach, 1912) Periods of cyclic ovarian activity in sexually mature females coincide with cycles of mammary growth, the duration of growth and the degree of glandular differentiation achieved being intimately related to the activity of the corpus luteum

The relationship between the mammae and the ovary was keenly appreciated by Thomas Beatson (1896) a Glasgow surgeon who sixty years ago performed oophorectomy in human breast cancer patients on the assumption that as the ovary controls the proliferation of normal mammary epithelium it may also control cellular proliferation in cancers arising from this epithelium He published the following conclusions in the *Transactions of the Medico Chirurgical Society of Glasgow* in 1910 I have never said that the procedure cures cancer, but I maintain and there is ample proof of it that in certain cases it certainly favourably controls the disease One of the values attached to oophorectomy is that the effects produced seem to me to have their chief interest and importance in that they throw a light upon the nature of carcinoma as a disease

In this way the general conception of hormone dependence in breast cancer came into being, the modern treatment of the disease by endocrine ablation is the direct though long delayed consequence of the work of this pioneer surgeon Eight years after the publication of Beatson s first paper, Hugh Lett (1905) published an analysis of ninety five cases of breast cancer treated by oophorectomy By this time it was fully realized that regression of the growth of the tumour and its metastases was invariably temporary We are still profoundly ignorant of the real nature of hormone dependence in cancers arising in organs which are themselves hormone dependent, but it is probably safe to assume in the case of hormone dependent breast cancer that the hormones essential for progressive proliferation of the cells of this tumour are those produced under physiological con ditions for normal mammogenesis It may be that these physiological hormones are produced by the cancer patient in excessive quantities, or in unphysiological proportions It may even be that the normal mammogenic hormones are changed or modified by the breast cancer patient, but there is literally no evidence either to support or to disprove these purely hypo thetical conceptions

It is abundantly clear however, that the intelligent investi gation of hormone dependence in cancer, as well as its rational treatment by endocrine ablation, must be based upon our know ledge of breast physiology and more especially of the complex interplay of those hormones which control the proliferation by mitosis of normal mammary epithelium I propose, therefore, to review briefly our present knowledge of the parts played by the ovaries, adrenal and hypophysis in the preliminary growth phases of mammogenesis which eventually lead to the forma tion of a functionally competent organ

Mammogenesis Induced by Ovarian Hormones in the Intact Animal

1 *Ovarian Oestrogenic Hormones*

The first demonstration that ovarian hormones stimulate the proliferation and differentiation of mammary epithelium was made by Vintemberger in 1925 Using ovarian follicular fluid

he produced a mammary growth response in rabbits which was confined to free production of new ducts without glandular differentiation or the formation of milk The first oestrogenic steroid hormone, oestrone was isolated in 1929 and from then until 1934 a series of observers, using purified oestrogens induced this limited response in intact rats male and female monkeys, dogs mice and cats On the other hand, between 1928 and 1940 a striking species difference in the response was demonstrated in intact guinea pigs virgin female goats and the virgin heifer, for in these animals oestrogen alone induced not only duct proliferation but widespread acinization and lactogenesis

It is now obvious that the mammae of all intact mammals are responsive to exogenous oestrogen, that both sexes respond and that mature, immature and weanling animals are equally responsive In the majority of mammals oestrogen alone induces growth of the duct system in the minority the gland becomes differentiated and lactogenesis is induced Another important general conclusion is that the optimal dose levels of oestrogen are low and lie within narrow limits, and there is general agreement that large doses are inhibitory This latter fact is of importance in the treatment of breast cancer by additive therapy

2 *Ovarian Progestational Hormones*

The administration to the intact mammal of progesterone alone has little effect on mammogenesis but progesterone together with oestrogenic steroids form a powerful mammogenic and lactogenic synergistic combination which brings the male or female mature or immature gland to the progestational state—in some cases to the structural status of mid pregnancy To produce an optimal response the two hormones must be administered in optimal proportions, an excess of either being inhibitory

THE ADRENAL GONAD RELATIONSHIP

Adrenal cortical tumours accompanied by virilism are known to produce excessive quantities of androgen those accompanied by feminism produce excessive amounts of oestrogen Following

In this way the general conception of hormone dependence in breast cancer came into being the modern treatment of the disease by endocrine ablation is the direct though long delayed consequence of the work of this pioneer surgeon Eight years after the publication of Beatson's first paper, Hugh Lett (1905) published an analysis of ninety five cases of breast cancer treated by oophorectomy By this time it was fully realized that regression of the growth of the tumour and its metastases was invariably temporary We are still profoundly ignorant of the real nature of hormone dependence in cancers arising in organs which are themselves hormone dependent, but it is probably safe to assume in the case of hormone dependent breast cancer that the hormones essential for progressive proliferation of the cells of this tumour are those produced under physiological conditions for normal mammogenesis It may be that these physiological hormones are produced by the cancer patient in excessive quantities, or in unphysiological proportions It may even be that the normal mammogenic hormones are changed or modified by the breast cancer patient, but there is literally no evidence either to support or to disprove these purely hypothetical conceptions

It is abundantly clear, however, that the intelligent investigation of hormone dependence in cancer, as well as its rational treatment by endocrine ablation, must be based upon our knowledge of breast physiology and more especially of the complex interplay of those hormones which control the proliferation by mitosis of normal mammary epithelium I propose therefore, to review briefly our present knowledge of the parts played by the ovaries adrenal and hypophysis in the preliminary growth phases of mammogenesis which eventually lead to the formation of a functionally competent organ

Mammogenesis Induced by Ovarian Hormones in the Intact Animal

1 *Ovarian Oestrogenic Hormones*

The first demonstration that ovarian hormones stimulate the proliferation and differentiation of mammary epithelium was made by Vintemberger in 1925 Using ovarian follicular fluid

he produced a mammary growth response in rabbits which was confined to free production of new ducts without glandular differentiation or the formation of milk The first oestrogenic steroid hormone, oestrone, was isolated in 1929 and from then until 1934 a series of observers using purified oestrogens in duced this limited response in intact rats male and female monkeys, dogs mice and cats On the other hand, between 1928 and 1940 a striking species difference in the response was de monstrated in intact guinea pigs virgin female goats and the virgin heifer for in these animals oestrogen alone induced not only duct proliferation but widespread acinization and lacto genesis

It is now obvious that the mammae of all intact mammals are responsive to exogenous oestrogen, that both sexes respond and that mature, immature and weanling animals are equally responsive In the majority of mammals oestrogen alone induces growth of the duct system, in the minority the gland becomes differentiated and lactogenesis is induced Another important general conclusion is that the optimal dose levels of oestrogen are low and lie within narrow limits and there is general agree ment that large doses are inhibitory This latter fact is of importance in the treatment of breast cancer by additive therapy

2 *Ovarian Progestational Hormones*

The administration to the intact mammal of progesterone alone has little effect on mammogenesis but progesterone together with oestrogenic steroids form a powerful mammogenic and lactogenic synergistic combination which brings the male or female mature or immature gland to the progestational state—in some cases to the structural status of mid pregnancy To produce an optimal response the two hormones must be administered in optimal proportions, an excess of either being inhibitory

THE ADRENAL GONAD RELATIONSHIP

Adrenal cortical tumours accompanied by virilism are known to produce excessive quantities of androgen those accompanied by feminism produce excessive amounts of oestrogen Following

the 'physiological castration' of the menopause, oestrogen production, after an initial fall, is frequently re established and can be again abolished, at least temporarily, by adrenalectomy This operation in ovariectomised or post menopausal women is not infrequently followed by re establishment of oestrogen production but the source of this oestrogen is unknown There is little doubt that loss or atrophy of the ovary and cessation of ovarian oestrogen and progesterone production *in women* stimulates the production of ovarian hormones by the adrenal cortex in gross excess of the low physiological output of the normal adrenal in the intact animal

The production of a gross excess of adrenal oestrogen in the ovariectomized animal has been studied in certain strains of inbred mice in which the adrenal 'take over is striking The results are inconstant if the ovaries are removed after the age of 3 months but if gonadectomized at the age of 3 days all animals show a vigorous response There is gross adrenal enlargement from diffuse or nodular hyperplasia or from the development of adenoma or carcinoma That this response is accompanied by excessive oestrogen production is shown by the development of cystic glandular hyperplasia of the endometrium, vaginal cornification and a vigorous growth reaction in the duct system of the mammary glands The growth response in the adrenal, whatever its nature is known to be mediated through the pituitary by its adreno cortical hormone and it can be abolished by exogenous oestrogen

The Hypophysis and Normal Mammogenesis

The study of growth responses in mammary tissue induced by exogenous steroid hormones in the intact animal leaves unanswered many fundamental questions regarding the overriding control exercised by the hypophysis in normal mammogenesis The story of pituitary participation began in 1928, when Stricker and Grueter administered unfractionated aqueous anterior pituitary extracts to ovariectomised pseudopregnant rabbits and clearly demonstrated that such extracts were capable of initiating lactation In the next year, Evans and Simpson demonstrated that the lactational response in the

rat was preceded and accompanied by an unequivocal growth response in the mammae, and this was confirmed in 1930 by Corner Three years later using fractionated pituitary extracts having high lactogenic potency, Lyons and Catchpole induced growth and acinization of the mammary ducts and initiated lactation

About 1930 a growing conviction arose that, to define the function of a target organ such as the ovary it was essential to use hypophysectomized animals and that the biological response to any substance deemed to be or to contain an anterior pituitary mammogenic hormone cannot be accepted as specific unless it can be repeatedly elicited without any doubt in the *non pregnant* hypophysectomized animal It is rather surprising that this conviction amongst workers engaged in the field of mammary physiology did not become apparent until the lapse of ten years after Philip E Smith, by establishing techniques for hypophysectomy in small mammals had demonstrated the relationship between the hypophysis on the one hand and somatic growth and sexual function on the other The reason for this delay was obviously the lack of any really conclusive experimental evidence that the hypophysis played any part in mammary growth and function

Between 1932 and 1935 six careful studies showed that the ovarian steroids alone could induce mammary growth in hypophysectomized animals This disquieting situation was resolved by Gomez and Turner in 1937 when they proved that if hypophysectomy is incomplete, ovarian steroids are mammogenic and the response produced may be even greater than in the intact animal They found that if the fragment left behind is only one fiftieth of the total weight of the anterior pituitary this anomalous response could be elicited It now became clear that no valid conclusions could be drawn from experiments using hypophysectomized animals unless careful scrutiny of the operation site aided if necessary by critical histological examination proves that pituitary ablation was total Since this time biochemical methods such as the estimation of protein bound iodine in the blood have become available and if the experiment can be carried out in young rapidly growing animals

failure to gain weight is a sound criterion of total ablation Following the establishment of this general principle, and between the years 1936 and 1937, it was well established that *the ovarian oestrogenic and progestational hormones acting alone or together are powerless to induce any mammary growth response in totally hypophysectomised mammals* This highly significant general conclusion was founded on experiments using rats, guinea pigs, mice, rabbits, cats and ground squirrels

In 1937, with these facts at their disposal, Gomez and Turner introduced their Pituitary Mammogen Hypothesis This stated that

1 A hitherto unrecognized anterior pituitary hormone or hormone complex is responsible for the growth and differentiation of the normal mammary gland

2 The pituitary mammogenic hormone or hormones are secreted in response to stimulation of the anterior pituitary by ovarian steroid hormones

It will be seen that according to this hypothesis the part played by the oestrogenic steroids in normal mammogenesis is reduced to a single function, i e the stimulation of the pituitary to produce a mammogenic hormone it does not permit the assumption that ovarian hormones have any other action in normal mammogenesis and it disallows the specific participation of any of the six well characterized hypophyseal hormones in this process

The mammogen hypothesis has been severely criticized but I will leave you to reconsider it when the experimental work conducted on this problem since 1937 has been reviewed Much of this work took advantage of the fact that fractionation of anterior pituitary extracts began to yield purified fractions having increasing lactogenic potency and increasing freedom from gross contamination with other well known hypophyseal hormones Such purified extracts as were available in 1941 and 1942 enabled Gardner and White to carry out a crucial experiment in totally hypophysectomized mice, the results of which are shown in Table 1

In 1943 these results were confirmed and amplified and by 1949 several claims had been made that lactogenic hormone,

now usually referred to as prolactin is a homogeneous protein, these claims being based on the appropriate physico-chemical criteria available at that time

In 1949 a significant observation was made by Desclin which suggested that, although the luteinizing fraction of pituitary

TABLE 1 Mammary growth response in hypophysectomized mice given various combinations of hypophyseal hormones—Gardner and White 1941 and 1942

Hormone	Mammary growth response	Duct growth	Acinization of ducts
Oestrogen and progesterone	Nil	Nil	Nil
Purified lactogenic hormone (prolactin)	±	±	Nil
Oestrogen progesterone + purified lactogenic hormone	++	++	++

gonadotrophin is responsible for luteinization of the ovum free follicle, the continued existence of the corpus luteum and its production of progesterone is controlled by pituitary prolactin He obtained a full luteotrophic response in the ovaries of intact virgin female rats by the administration of prolactin and this response was accompanied by extensive mammary growth and the formation of uterine deciduomata During the next few years the results of further experimental work on hypophysectomized rats—to be referred to later—confirmed, in the rat at least that prolactin is the trophic hormone responsible for the maintenance of the corpus luteum and therefore, for the production and release of progesterone

A long and elaborate series of crucial experiments carried out by Lyons Li and their associates between 1950 and 1955 clarified the situation in many directions Table 2 is a brief summary of their first major contribution

These results show clearly that three hormones—oestrogen, progesterone and prolactin—are essential for normal mammogenesis No single member or any pair of hormones in this triad can specifically stimulate mammary growth in the absence of the pituitary and gonads, and any combination of hormones

which does not include prolactin is equally powerless to do so It is usually assumed that the action of growth hormone and A C T H , by helping to restore the hypophysectomized animal to a more physiological state, brings the mammary gland re acting to the hormones of the essential mammogenic triad to a higher degree of functional competence

TABLE 2 Mammogenesis in hypophysectomized and oophorectomised rats experiments of Lyons Li and associates 1950–2

Given	Result	
1 Oestrone	Mammary atrophy	Pituitary mammotrophin lacking
2 Oestrone and progesterone	Mammary atrophy	Pituitary mammotrophin lacking
3 Purified pro lactin	Regression delayed No mammary growth	Ovarian hormones lacking
4 Purified pro-lactin + oestrone	Regression delayed No mammary growth	Progesterone lacking no ovary
5 Oestrone + pro gesterone + puri fied prolactin	Mammogenesis + + to status of mid pregnancy No lactogenesis	Full lactogenesis + lactopoiesis needs adrenal cortical hor mones (*vide infra*)
6 Oestrone + pro gesterone + crude prolactin	Mammogenesis + + + to full term pregnancy Lactogenesis ±	Crude prolactin may contain A C T H or G H or both
7 Oestrone + pro gesterone + puri fied prolactin + G H	As above but lac togenesis +	
8 Oestrone + pro gesterone + puri fied prolactin G H + A C T H	Growth to full term pregnancy Lactogenesis + + +	Full lactogenesis needs intact pituitary adrenal axis

In 1950 the same investigators succeeded in completely fem inizing the mammary glands of hypophysectomized im mature male rats The mammae were again brought to the structural status of mid pregnancy by the mammogenic triad, to the structural and functional status of full term pregnancy by the addition of growth hormone and cortisone, and lactation was initiated in these previously rudimentary male glands In all these experiments the induction of mammogenesis depended



FIG 5 *Photographs of whole gland mounts* (× 6 7) Mammae dissected out of defatted *carmine stained pelt and mounted* without sectioning
Row 1 *Slowly growing* unclubbed mammae of intact untreated male weanling
Row 2 *Atrophic* mammae produced by total hypophysectomy Neither ovarian hormones nor prolactin when given alone can retard this process
Rows 3 and 4 *Rapidly growing* heavily clubbed mammae of totally hypophysectomized male weanling mice reacting to the mammogenic triad
Row 3 evolution of the reaction Row 4 maximal response after five days treatment The gland on the right is heavily acinized
Row 5 *Rapidly growing* mammae of totally hypophysectomized animals reacting to the administration of a combination of an extract of human female urine and the ovarian hormones oestrone and progesterone

upon the use of prolactin, oestrone and progesterone in optimal proportions the dose of oestrone being remarkably low whilst those of progesterone and prolactin were correspondingly high in all investigations of this kind it is obvious that careful *in vivo* titration of the hormones used is an essential preliminary to the main experiment

During the same time, Nelson, using another approach, confirmed the general conclusions reached by Lyons and his colleagues and added further evidence in favour of the luteotrophic function of prolactin He administered oestrogen at a high dose level to rats for 8 to 20 days, they became pseudopregnant developed large corpora lutea and an unmistakable growth response was induced in their mammary glands It seems safe to assume that in this experiment the exogenous oestrogen, by stimulating the hypophysis was responsible for the release of prolactin and that this hormone, by virtue of its luteotrophic action, was responsible for the enlargement of the corpora lutea and the release of progesterone The mammary response was presumably induced by the combination of exogenous oestrogen and endogenous prolactin and progesterone In another experiment the period of adminstration of oestrogen was shortened to 8 to 14 days and the animals were then hypophysectomized The reacting mammary glands became atrophic and the corpora lutea were no longer maintained Prolactin was then administered and unmistakable stimulation of mammary growth and differentiation together with reactivation of the atrophic corpora lutea followed

There is enough direct and indirect evidence available to be reasonably confident that the release of prolactin from the hypophysis is due to hypophyseal stimulation by circulating oestrogen and that this stimulus is only effective at relatively low concentrations of oestrogen (Figure 1) Above a low threshold the stimulus becomes increasingly ineffective and at a high threshold inhibitory This fundamental mechanism is clearly responsible not only for providing prolactin for normal mammogenesis but also for the maintenance of the corpus luteum and the production and release of progesterone (see Figures 2, 3 and 4)

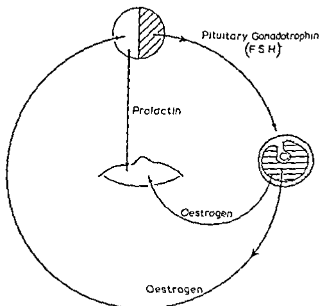


FIG 1 Prolactin release by oestrogen

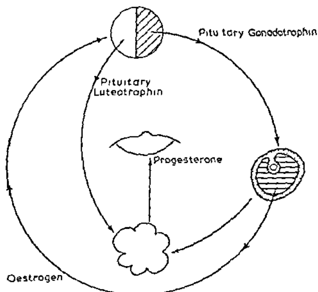
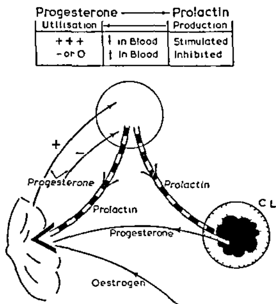


FIG 2 Release of progesterone by pituitary luteotrophin (? prolactin)



Progesterone Utilisation	←	→ Prolactin Production
+ + +	↓ in Blood	Stimulated
– or O	↑ in Blood	Inhibited

Fig 3 The corpus luteum—one of the prolactin target organs

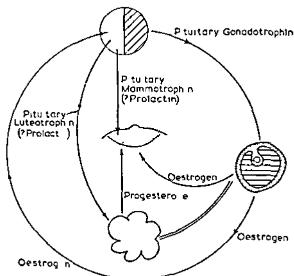


Fig 4 Correlation of production and utilization of the mammotrophic pituitary triad

Does prolactin exert a direct and specific effect on mammary epithelium? An experiment performed by Lyons in 1942 suggests that it does He injected purified prolactin directly into single galactophores of the rabbit mamma Lactation was initiated in the injected sectors of the gland No secretory response was observed in uninjected sectors Histological examination of the stimulated sectors showed a significant increase in the number of mitotic figures in the mammary epithelium The experiment would have provided more solid evidence if it had been performed on hypophysectomized animals

There have been repeated claims, over the years, that the local and specific action of oestrogenic steroids on susceptible tissues is characterized by capillary hyperaemia, dilatation and increased permeability It is not unreasonable to suppose, therefore, that by inducing capillary hyperaemia in the mammary gland, the escape of the relatively small molecule of prolactin into the tissue fluid of the breast may be materially facilitated *It remains to be proved that oestrogens play a part* in stimulating the process of mitosis in mammary epithelium as they do in other tissues It is not improbable that they are responsible for the mesenchymal proliferation in the mammary stroma which eventually provides a collagen framework for the gland

The probable sequence of events in normal mammogenesis may be briefly summarized as follows

1 The production and release of ovarian and adrenal oestrogen effected by pituitary gonadotrophin

2 The release of prolactin effected by oestrogenic stimulation, Figure 1

3 Prolactin exerts its three mammotrophic activities

(*a*) As a *luteotrophin* it controls the production and release of progesterone Figures 2 and 3

(*b*) As a *mammotroph n* acting in synergism with progesterone and oestrogen it is responsible for the growth and differentiation of mammary epithelium Figure 4

(*c*) As a *lactogenic* hormone it initiates lactation and in synergism with other hormones it is concerned in lactopoiesis

Application to Breast Cancer

The results of surgical ablation of endocrine organs in patients suffering from breast cancer have taught us to expect a significant and worthwhile growth regression in the tumour and its metastases in approximately 40 per cent of patients This figure would probably be higher if the technique for hypophysectomy could be guaranteed to remove all pituitary tissue Ablation operations however, add little to our knowledge of the true nature of either hormone dependence or independence, and it would seem that repeated determinations on a research basis, of the hormonal status of the breast cancer patient—studied in parallel with the clinical course of the disease—in a sufficiently large series of patients is very likely to be of material assistance in the solution of this complex problem It is also probable that such studies would help in the selection of patients likely to benefit by each type of ablation operation An inquiry such as this presents many administrative and scientific difficulties Some of these difficulties have been surmounted over the last few years in the Clinicopathological Laboratories of the Imperial Cancer Research Fund where the excretion of the phenolic oestrogenic steroids has now been estimated in a large series of patients before and after each type of endocrine ablation (Young Bulbrook and Greenwood, 1957, Bulbrook and Greenwood 1957, Greenwood and Bulbrook 1957) The methods used are based on those proposed by Brown in 1955 The urine is subjected to acid hydrolysis and extracted with ether, the extract is dissolved in alcohol, benzene and petroleum ether are added and the product is extracted with water The water extract carries the oestriol fraction This is methylated, and caustic soda and hydrogen peroxide added A final extraction with ether yields a product which is adsorbed on an alumina column and estimated by the Kober reaction as oestriol methyl ether Similar treatment of the water insoluble primary ether extract yields oestrone and oestradiol 17β methyl ethers which are estimated by the same reaction This accurate method of estimation is expensive and time consuming and being unsuitable for routine use can only be undertaken

on a 'research' basis Pregnanediol, the metabolic end product of progesterone, can be estimated by a method which has the same clinical limitations

The investigation of pituitary function in the breast cancer

TABLE 3 Mammary growth responses in hypophysectomized male weanling mice (A_2G strain) showing that human female urine is an effective substitute for prolactin in the Mammogenic Triad (prolactin oestrone and progesterone)

Group	Hormone administered	Result
1	None	Mammary atrophy obvious on 5th day well established on 10th day
2	Oestrone and progesterone	Mammary atrophy
3	Prolactin	Mammary atrophy
4	Oestrone progesterone and prolactin	A vigorous growth response in almost all mammary glands of all animals
5	Oestrone and progesterone and human female urine	Growth response in almost all mammae of all animals identical with that in Group 4 but rather less vigorous

patient is a problem of considerable difficulty For the estimation of pituitary gonadotrophin a standard bio-assay is available A method which promises to yield information regarding the output of pituitary mammotrophin in the urine is being investigated in the Clinicopathological Laboratories of the Imperial Cancer Research Fund It is based on the demonstration that the urine of all normal pre menopausal women in the second half of the menstrual cycle, when injected into intact weanling male mice, will induce an easily recognisable growth reaction in their rudimentary mammary glands (Scowen and Hadfield, 1955, Hadfield and Young, 1956a) The urine of 57 per cent of post menopausal women contains the same mammogenic agent (Hadfield and Young, 1956b) The capacity to react to this urinary mammotrophin is, unfortunately, confined to a small minority of all easily available strains of laboratory mice (Young, 1957) This entails breeding and maintaining the susceptible strain and detracts from the value of the test as a routine procedure

In an attempt to define the nature and origin of this urinary mammotrophic agent a series of experiments has been carried out on totally hypophysectomized weanling male mice (Hadfield, 1957) A control group of animals was left untreated and killed on the tenth post operative day In other groups the male weanlings were injected with various hormone combinations for five days, the first dose in each case being given on the fifth day after hypophysectomy These animals were also killed on the tenth post operative day The results are given in Table 3 and can be seen in Plate XXIII Figure 5 [1]

These experiments show that normal female urine contains a hormone which is either pituitary prolactin or has a specific biological action identical with that of prolactin They also promise to provide a method for the estimation of mammotrophic function in the breast cancer patient By using hypophysectomized mice the difficulty associated with strain variation in the mammary response to the injection of urine will, in all probability, disappear

[1] This plate will be found facing page 304

REFERENCES

BEATSON G T (1896) *Lancet* ii 104
BEATSON G T (1910) *Trans Med Chir Soc Glasgow* 9 166
BROWN J B (1955) *Biochem J* 60 185
BULBROOK R D and GREENWOOD F C (1957) *Brit med J* i 662
CORNER G W (1930) *Amer J Physiol* 95 43
DESCLIN L (1949) *Ciba Colloq Endocrinol* 4 395
EVANS H M and SIMPSON M E (1929) *Proc Soc exper Biol Med* 26 597
GARDNER W U and WHITE A (1941) *Proc Soc exp Biol N Y* 48 590
GARDNER W U and WHITE A (1942) *Anat Rec* 82 414
GOMEZ E T and TURNER G W (1937) *Missouri Agric Exp Station Res Bull* No 259
GREENWOOD F C and BULBROOK R D (1957) *Brit med J* i 666
GREGORIEV W (1897) *Zbl Gynak* 21 663
HADFIELD G (1957) *Lancet* i 1058
HADFIELD G and YOUNG J S (1956a) *Brit J Cancer* 10 145
HADFIELD G and YOUNG J S (1956b) *Brit J Cancer* 10 324
LETT H (1905) *Trans Roy Med and Chir Soc London* 88 147

on a 'research' basis Pregnanediol, the metabolic end product of progesterone, can be estimated by a method which has the same clinical limitations

The investigation of pituitary function in the breast cancer

TABLE 3 Mammary growth responses in hypophysectomized male weanling mice (A_4G strain) showing that human female urine is an effective substitute for prolactin in the Mammogenic Triad (prolactin oestrone and progesterone)

Group	Hormone administered	Result
1	None	Mammary atrophy obvious on 5th day well established on 10th day
2	Oestrone and progesterone	Mammary atrophy
3	Prolactin	Mammary atrophy
4	Oestrone progesterone and prolactin	A vigorous growth response in almost all mammary glands of all animals
5	Oestrone and progesterone and human female urine	Growth response in almost all mammae of all animals identical with that in Group 4 but rather less vigorous

patient is a problem of considerable difficulty For the estimation of pituitary gonadotrophin a standard bio-assay is available A method which promises to yield information regarding the output of pituitary mammotrophin in the urine is being investigated in the Clinicopathological Laboratories of the Imperial Cancer Research Fund It is based on the demonstration that the urine of all normal pre menopausal women in the second half of the menstrual cycle, when injected into intact weanling male mice will induce an easily recognisable growth reaction in their rudimentary mammary glands (Scowen and Hadfield, 1955, Hadfield and Young, 1956a) The urine of 57 per cent of post menopausal women contains the same mammogenic agent (Hadfield and Young 1956b) The capacity to react to this urinary mammotrophin is unfortunately confined to a small minority of all easily available strains of laboratory mice (Young, 1957) This entails breeding and maintaining the susceptible strain and detracts from the value of the test as a routine procedure

XVIII

The Biliary Excretion of Thyroid Hormone

N B MYANT

THE importance of the liver in the metabolism of organic iodine was first recognized by Kendall (1919) He injected 200 mgm of thyroxine intravenously into a dog and later recovered more than 40 per cent of the injected iodine in the bile In this experiment he was unable to establish the chemical form of the biliary iodine However when radioactive thyroxine was made available (Joliot Courrier, Horeau and Sue 1944) it became possible to study the biliary excretion of thyroxine under more physiological conditions and to analyse the chemical form of the iodine appearing in the bile

Gross and Leblond (1950) injected radioactive thyroxine into rat ands showed that it was removed from the body in two ways Some of the thyroxine was broken down in the tissues and the radioactive iodine excreted in the urine as iodide but a large proportion of the dose was excreted in the bile and faeces as organic iodine They also showed that the proportion of the dose excreted in the faeces was related to the amount of thyroxine injected When they gave a small dose most of the radio-iodine was excreted in the urine only about 20 per cent appearing in the faeces But when they gave a large dose more than 80 per cent was excreted in the faeces They suggested that the biliary pathway of excretion might act as an overflow for removing an excess of thyroid hormone from the body By similar methods Albert and Keating (1952) also showed that exogenous thyroxine is rapidly removed by the liver and excreted in the bile

LYONS W R (1942) *Proc Soc exp Biol N Y* **51** 308

LYONS W R and CATCHPOLE, H R (1933) *Proc Soc exp Biol N Y* **31** 299

LYONS W R JOHNSON R E COLE R D and LI C G (1955) *Hypophyseal Growth Hormone Nature and Actions* International Symposium New York (McGraw Hill Book Co Inc)

NELSON W O (1952) *Ciba Colloq Endocrinol* **4** 402

SCOWEN E F and HADFIELD G (1955) *Cancer* **8** 890

SMITH P E (1923) *Endocrinol* **7** 579

STEINACH E (1912) *Pflug Arch ges Physiol* **144** 71

STRICKER P and GRUETER F (1928) *C.R Soc Biol Paris* **99** 1978

VINTEMBERGER P (1925) *Arch Biol (Paris)* **35** 125

YOUNG S (1957) *Brit J Cancer* In the press

YOUNG S BULBROOK R D and GREENWOOD F C (1957) *Lancet* **ii** 350

glucuronides formed in the body As I shall explain below, work on the biosynthesis of glucuronides of *substituted phenols* suggests that thyroxine glucuronide is synthesized through a rather complicated pathway

The formation of thyroxine glucuronide in the rat's liver also occurs in physiological conditions when no exogenous

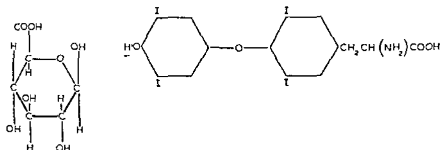


FIG 2 Probable structure of Compound U

thyroxine is added to the body The endogenous thyroid hormone in the circulation can be labelled by injecting radioiodide into the animal a day or two before the bile samples are taken This allows time for the radioiodide to be taken up by the thyroid and returned to the plasma as radioactive hormone Radioactive thyroxine glucuronide can then be shown to be present in the bile (Taurog *et al* 1952) Indeed, the proportion of the total ^{131}I in the bile excreted as the glucuronide tends to be greater in these conditions than when exogenous radioactive thyroxine is injected into the rats

In addition to the two bands of blackening corresponding to thyroxine and thyroxine glucuronide (Figure 1) two other bands are sometimes seen One of these about halfway between thyroxine and thyroxine glucuronide, has been shown by Roche Michel and Tata (1954) to be due to a substance moving with the same R_F values as the pyruvic acid analogue of thyroxine—that is thyroxine in which the alanine side chain has been oxidized to pyruvic acid The other band just visible in Figure 1, lies between the origin and thyroxine glucuronide Nothing is known about the substance that gives rise to this band

Chemical Form of the Iodine in the Bile

With regard to the chemical form of the biliary iodine, the first step was taken by Taurog, Briggs and Chaikoff in 1951 They injected radioactive thyroxine into rats and then analysed the radioactive substances in the bile by means of paper chromatography The positions to which the various radioactive substances moved were found by making autoradiographs Plate XXIV, Figure 1 shows two of their autoradiographs[1] Furthest from the origin there is a band of blackening corresponding to the position of free thyroxine About halfway between thyroxine and the origin is another band The substance giving rise to this band they called Compound U When large amounts of thyroxine were given as carrier with the injection, the band at thyroxine was more intense than the one at Compound U (Figure 1b), but with small amounts of carrier, Compound U contained nearly all the radioactivity on the chromatogram (Figure 1a) Evidently, there is a limit to the capacity of the liver for converting thyroxine to Compound U

By eluting Compound U from the paper chromatogram Taurog *et al* (1952) were able to study some of its properties Free thyroxine could be released from it by alkaline hydrolysis so it must be a conjugate of thyroxine Thyroxine was also released from Compound U by the action of β glucuronidase Since this enzyme has a high degree of specificity towards its substrate, Compound U is probably a glucuronide of thyroxine Although this cannot yet be regarded as certain, for convenience I shall refer to it as thyroxine glucuronide' Since Klitgaard, Lipner, Barker and Winnick (1953) have shown that the amino and carboxyl groups of thyroxine are free in Compound U, the glucuronic acid is probably attached to the end of the thyroxine molecule opposite to the alanine side chain If so, the structure of the glucuronide would be that shown in Figure 2 Figure 2 only shows the structural relations of the two parts of the glucuronide molecule and is not meant to imply that it arises by a condensation between free glucuronic acid and free thyroxine In fact free glucuronic acid is not a precursor of

[1] The plates referred to in this lecture will be found between pages 3 0-1

The Enterohepatic Circulation of Thyroid Hormone

The more physiological aspects of the biliary excretion of thyroxine glucuronide have also been investigated by Taurog and his colleagues Radioactive glucuronide does not appear in the plasma of rats after injections of radioactive thyroxine However, if radioactive thyroxine glucuronide is eluted from a chromatogram of ^{131}I labelled bile and is then put into a rat s intestine free radioactive thyroxine appears in the faeces This shows that the glucuronide excreted in the bile is to some extent hydrolysed by the β glucuronidase present in the intestine Since free thyroxine is readily absorbed from the intestine, some of the thyroxine excreted in the bile as the glucuronide must ultimately be reabsorbed into the blood as free thyroxine Roche, Michel Michel and Tata (1954) have shown that tri iodothyronine undergoes a similar cycle of conjugation biliary excretion and reabsorption Thyroxine and tri iodothyronine together account for the circulating form of the thyroid hormone, so the excretion and reabsorption of these two substances may be described as an enterohepatic circulation of thyroid hormone

This brings us to the question How extensive is the biliary excretion and reabsorption of thyroid hormone? The complete answer to this would tell us how many micrograms of hormone are excreted in the bile each day how many are reabsorbed and how many are lost in the faeces We should also like to know how these quantities compare with the daily rate of synthesis of hormone by the thyroid

Biliary and Faecal Clearance Rates of Thyroid Hormone

It is most convenient to measure these quantities as clearance rates By analogy with the renal clearance rate of urea, the rate at which any organ excretes a substance from the plasma can be expressed as a volume of plasma cleared in unit time In general the clearance rate of any plasma constituent is equal to

$$\frac{C_f \times V}{C_p}$$

There is some evidence that all these derivatives of thyroxine are excreted in human bile During operations to the upper abdomen it is sometimes possible to obtain bile from the gall bladder of a person with a normal liver If the patient has been given an injection of radioactive thyroxine some hours before

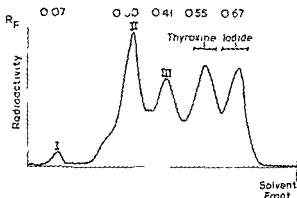


FIG 3 Radioactivity along chromatogram made from bile taken from the gall bladder of a patient given an injection of ^{131}I labelled l thyroxine R_F values of peaks and markers as shown Solvents collidine and water

the operation the iodine in his bile can be analysed by paper chromatography Figure 3 shows the results obtained in one patient As there was not enough radioactivity on the chromatogram to give clearly visible blackening on an autoradiograph, the activity along the paper was measured with a continuously recording scanner In this patient and in the few others who were studied (Myant 1956b), three peaks of radioactivity were observed on the chromatogram All the three peaks (Figure 3 I, II and III) lie behind the position to which free thyroxine moves, there being no peak at either the thyroxine or the iodide marker From this result alone, the substances responsible for these peaks cannot be identified However peak II has about the same R_F value as thryoxine glucuronide and peak III has about the same R_F value as the pyruvic acid analogue of thyroxine These results are at least consistent with the supposition that in humans the metabolism of thyroxine in the liver is the same as it is in rats The only difference seems to be that in human bile there is very little free thyroxine

The Enterohepatic Circulation of Thyroid Hormone

The more physiological aspects of the biliary excretion of thyroxine glucuronide have also been investigated by Taurog and his colleagues Radioactive glucuronide does not appear in the plasma of rats after injections of radioactive thyroxine However if radioactive thyroxine glucuronide is eluted from a chromatogram of ^{131}I labelled bile and is then put into a rat s intestine free radioactive thyroxine appears in the faeces This shows that the glucuronide excreted in the bile is to some extent hydrolysed by the β glucuronidase present in the intestine Since free thyroxine is readily absorbed from the intestine some of the thyroxine excreted in the bile as the glucu ronide must ultimately be reabsorbed into the blood as free thyroxine Roche, Michel Michel and Tata (1954) have shown that tri iodothyronine undergoes a similar cycle of con jugation, biliary excretion and reabsorption Thyroxine and tri iodothyronine together account for the circulating form of the thyroid hormone so the excretion and reabsorption of these two substances may be described as an enterohepatic circula tion of thyroid hormone

This brings us to the question How extensive is the biliary excretion and reabsorption of thyroid hormone? The complete answer to this would tell us how many micrograms of hormone are excreted in the bile each day how many are reabsorbed and how many are lost in the faeces We should also like to know how these quantities compare with the daily rate of syn thesis of hormone by the thyroid

Biliary and Faecal Clearance Rates of Thyroid Hormone

It is most convenient to measure these quantities as clearance rates By analogy with the renal clearance rate of urea the rate at which any organ excretes a substance from the plasma can be expressed as a volume of plasma cleared in unit time In general the clearance rate of any plasma constituent is equal to

$$\frac{C_r \times V}{C_{\cdot}}$$

where C_f is the concentration in the fluid excreted, C_p is the concentration in the plasma, and V is the volume of fluid excreted per hour This way of measuring the rate of excretion has the advantage that it may not be necessary to measure absolute concentrations, which may be difficult if, as in the case of thyroxine, the plasma concentration of the substance is very low If the thyroxine in the circulation is labelled with radioactive iodine, the rate at which the liver clears thyroxine from the plasma and excretes it into the bile (either as free thyroxine or as its breakdown or conjugation products) may be expressed as

$$\frac{(^{131}\text{I concentration in bile})}{(^{131}\text{I concentration in plasma})} \times (\text{rate of flow of bile})$$

This we may call the 'biliary clearance rate' of thyroxine

Similarly, the rate of excretion in the faeces may be expressed as the faecal clearance rate,

$$\frac{Q_f}{C_p}$$

where Q_f is the quantity of ^{131}I excreted in the faeces per hour and C_p is the concentration in the plasma In this case, the total amount of ^{131}I excreted is measured directly and is not, as in the case of the bile, obtained as the product of the concentration and the volume

Figure 4 shows the results of an experiment in which the biliary clearance rate of thyroxine was measured in a rat The rat was given an intravenous injection of 1 microgram of ^{131}I labelled 1 thyroxine and then, at zero time the bile duct was cannulated The bile was collected for an hour and the ^{131}I concentration measured in samples taken over ten minute intervals At the same time the plasma ^{131}I concentration was measured in serial blood samples The bile flow averaged about 0 7 ml per hour, and the bile/plasma ^{131}I concentration ratio (B/P ratio) varied from 2 to 3 showing a tendency to fall at the end of the experiment Since the biliary clearance rate is given by the product of the B/P ratio and the rate of flow of bile the clearance rate in this rat was about 2 ml of plasma per hour

The biliary clearance rate of endogenous thyroid hormone can be measured in the same way if the circulating hormone is labelled by giving the rat an injection of radioiodide 24 hours before cannulation of the bile duct In these conditions the B/P ratio lies between 2 and 3 and the bile flow is a little under

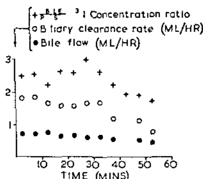


FIG 4 Bile/plasma ^{131}I concentration ratio bile flow and biliary clearance rate after injection of 1 microgram of ^{131}I labelled l thyroxine into a rat

1 ml per hour The clearance rate of endogenous hormone is, therefore about the same as the clearance rate of a dose of 1 microgram of exogenous thyroxine

The faecal clearance rate of endogenous hormone can be measured in rats prepared in the same way Figure 5 shows the daily rate of excretion of ^{131}I in the faeces of a rat in which the endogenous thyroid hormone in the circulation was labelled by ^{131}I The concentration of protein bound ^{131}I in the plasma is shown in the same figure The faecal clearance rate expressed in ml of plasma cleared per day is obtained by dividing the rate of excretion by the plasma concentration In this rat (Figure 6) the faecal clearance rate averaged about 20 ml of plasma per 24 hours At the end of the experiment the bile duct was cannulated and the biliary clearance rate measured If the biliary clearance rate is assumed to remain constant throughout the 24 hours the volume of plasma cleared into the bile each day can be estimated from the rate measured over a short interval In this case (Figure 6) the estimated value was 35 ml of plasma

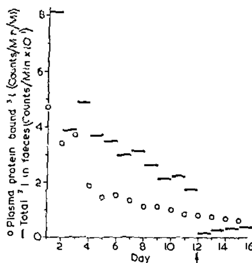


Fig 5 Plasma protein bound ^{131}I concentration (○) and faecal excretion of ^{131}I (—) in a rat in which endogenous hormone was radioactive before and after ligation of the bile-duct (shown by arrow)

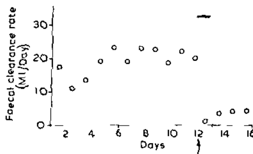


Fig 6 Faecal clearance rate of endogenous thyroid hormone (○) with single estimate of biliary clearance rate (ml /day) Values calculated from observations shown in Fig 5

PLATE XXIV

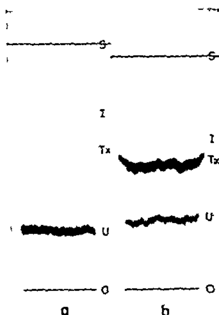


FIG 1 Radioautographs of *chromatograms* from bile of rats taken 3 to 4 hours after injection of ^{131}I labelled L-thyroxine (*a*) after 71 micrograms (*b*) after 5 0 micrograms Solvents collidine and water
(Reproduced with permission from Taurog Briggs and Chaikoff 1952)

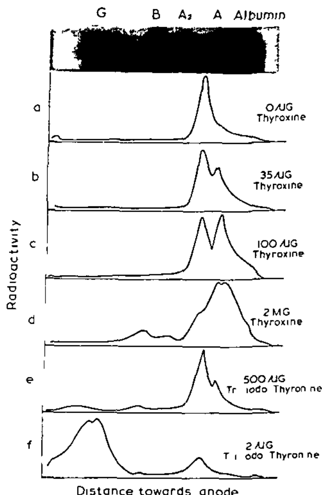


FIG 10 Paper electrophoresis of rat s serum in borate buffer at pH 7 ɔ with curves showing radioactivity along the paper *a–e* serum from rats with radioactive endogenous hormone *a* with no injection of thyroxine *b* with 35 µg *c* with 100 µg *d* with 2 mg thyroxine *e* with 500 µg tri iodothyronine *f* serum from normal rat after injection of 2 µg radioactive tri iodothyronine vertical scale adjusted to make the tallest peaks approximately the same height

We can now write down a balance sheet showing the relative amounts of thyroid hormone excreted in the bile reabsorbed into the circulation, and lost in the faeces Table 1 shows the averages of values derived from several rats The equivalent of about 45 ml of plasma are cleared into the bile each day, and

TABLE 1 Clearance of thyroid hormone in 24 hours (expressed in ml of plasma)

	Biliary	Faecal	Reabsorbed	Micrograms synthesized per day
Rats	+45 (2¼)	+20 (1)	−25	5
Humans	+500 (25)	+200 (10)	−300	150–200

Numbers in brackets give micrograms of hormone in the given volume of plasma

about 20 ml are cleared into the faeces The difference between the biliary and faecal clearance rates—about 25 ml—gives a measure of the amount reabsorbed from the intestine each day This is shown as a negative clearance rate since the iodine is moving back into the plasma Some of the iodine reabsorbed is presumably in the form of free thyroxine, but some may also be in the form of iodide or other iodine containing products of the metabolism of thyroid hormone The amount of organic iodine in 25 ml of plasma represents, therefore, the maximum amount of hormone that the rat can recover by reabsorption each day

These volumes of plasma cleared can be translated into micro grams of hormone excreted, by multiplying the clearance rates by the concentration of hormone in the plasma In a normal rat 100 ml of plasma contain about 5 micrograms of thyroid hormone Therefore the 45 ml of plasma cleared into the bile each day correspond to the excretion of 2¼ micrograms of hor mone (shown within brackets in Table 1) Although up to 50 per cent of this may be reabsorbed from the intestine there is a net loss of iodine in the faeces equivalent to about 1 microgram of hormone, i e the amount contained in 20 ml of plasma (Table 1) Compared with the amount of hormone synthesized by the rat s thyroid these quantities are considerable Normally

the day and night In round figures, the biliary clearance rate is about 500 ml of plasma in 24 hours Since the plasma con centration of thyroid hormone in a normal person is about 5 micrograms per 100 ml, this means that 25 micrograms of hormone, or an equivalent amount of iodine, are excreted in the bile in 24 hours Of this, about 10 micrograms (correspond ing to 200 ml of plasma) are lost in the faeces

The thyroid of a normal adult makes about 200 micrograms of hormone in 24 hours According to these results, therefore, a much smaller proportion of the day s output of hormone is excreted in human bile than in rat s bile Nearly half the output of hormone from a rat s thyroid is excreted in the bile (see above), whereas in the patients on whom these measurements were made, only about 25 micrograms out of every 200 syn thesized were excreted by this route However, the difference between humans and rats may not be quite as great as Table 1 suggests, because all the figures for the biliary clearance rate in humans are derived from patients who had recently had some obstruction to the flow of bile, and in whom the bile was pre vented from flowing into the duodenum The secretion of bile is known to be stimulated by bile salts reabsorbed from the duo- denum

RELATION BETWEEN BILIARY CLEARANCE RATE AND DOSE INJECTED I want now to return to a point I mentioned earlier on It will be remembered that Gross and Leblond found that if they in jected a large dose of thyroxine into a rat a greater proportion of the dose was excreted in the faeces than if they injected a small dose In trying to explain this, we might consider the possibility that the biliary clearance rate is higher with a large dose of thyroxine than with a small one The fact that an increase in the plasma concentration of the thyroxine sometimes leads to an increase in the biliary clearance rate in humans (Figure 7) could be taken as evidence for this But we can find much clearer evidence in rats, where the conditions can be varied more widely If rats are given injections of radioactive thyroxine con taining varying amounts of non radioactive thyroxine, the B/P ratio is found to increase as the amount of carrier is increased

Figure 8 shows the results of several experiments, each done on a separate rat, in which a small dose of radioactive thyroxine was injected followed by a second injection containing 10 to 2,000 micrograms of carrier With the largest dose of carrier the B/P ratio rose to nearly 50 Since the bile flow (not shown in

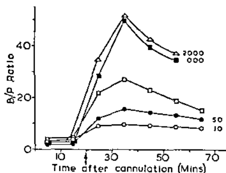


FIG 8 Bile/plasma ^{131}I concentration ratio after injection of 1 microgram of radioactive thyroxine into a rat showing effect of an injection of non radioactive thyroxine (shown by arrow) 20 minutes after cannulation Number in small figures micrograms of thyroxine in second injection

the figure) remained unchanged during these experiments the biliary clearance rate rose in proportion to the increase in the B/P ratio If 0 7 ml per hour is taken as an average value for the rate of flow of bile, the highest point on Figure 8 would represent a clearance rate of more than 35 ml per hour—at least three times a rat s plasma volume cleared per hour

The same result is obtained if non radioactive thyroxine is injected into rats while the clearance rate of endogenous hormone is measured Figure 9 shows the results of an experiment in which 50 micrograms of thyroxine were injected into a rat prepared by an injection of radioiodide 24 hours earlier Within 20 minutes of the injection of thyroxine the B/P ratio rose to more than 20 The B/P ratio rises to about the same value after a given amount of thyroxine whether it is injected into normal rats as radioactive exogenous thyroxine, or as non radioactive thyroxine into rats in which the endogenous hormone is labelled This is simply a consequence of the fact that the endogenous hormone in the circulation exchanges with injected thyroxine

In the one case the plasma thyroxine is labelled by radioactive thyroxine molecules introduced from outside, in the other, the thyroxine introduced from outside is labelled by radioactive thyroxine already in the circulation

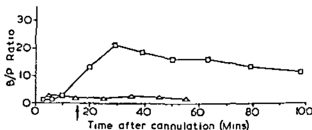


FIG 9 Bile/plasma concentration ratio of radioactive endogenous hormone △ without injection of carrier □ with an injection of 50 micrograms of non radioactive thyroxine (indicated by arrow)

One other fact about the biliary clearance rate may be mentioned, since, as I shall show later it helps to explain how the plasma concentration of thyroxine influences the biliary clearance rate An injection of tri iodothyronine has very little effect on the biliary clearance rate of endogenous hormone If, for instance, in an experiment similar to that in Figure 9 a rat is given an injection of 50 micrograms of non radioactive tri iodothyronine, the clearance rate scarcely rises above the resting value of 2 ml per hour

Relation between Biliary Clearance Rate of Thyroxine and its Binding by the Serum Proteins

It is helpful to consider the clearance rate as a measure of the efficiency of extraction—to borrow a term from renal physiology The volume of plasma cleared in unit time, expressed as a fraction of the plasma flow to the liver is a measure of the proportion of the thyroxine in the plasma that is extracted during its passage through the liver Now the plasma flow to a rat's liver is about 300 ml /per hour When, therefore the biliary clearance rate is 2 ml per hour the liver is extracting 2 out of every 300 molecules of thyroxine brought to it by the blood Provided the blood flow to the liver remains constant, the rise in clearance rate with increasing doses of thyroxine means that

the efficiency of extraction increases as the load to the liver is increased With the highest doses, the efficiency of extraction rises to about 12 per cent (35 ml of plasma cleared out of each 300 ml brought to the liver) This relationship is rather surpris ing, and is the opposite of the way in which the kidney behaves towards substances such as diodrast that are secreted by the tu bules With these substances, as the load to the kidney is increased a smaller proportion is extracted and so the clearance rate falls

A possible reason for this relationship is suggested by a con sideration of the way in which thyroid hormone is carried in the plasma

It was shown independently by Gordon, Gross, O Connor and Pitt Rivers (1952) and by Larson, Deiss and Albright (1952) that the circulating thyroid hormone in humans is bound to a specific fraction of the serum proteins This protein (thy roxine binding protein or TBP) can be separated by paper electrophoresis in a barbiturate buffer at pH 8 4 and is found to move to a position between the α_1 and α_2 serum proteins If the thyroxine in the serum is labelled with ^{131}I and the serum pro-teins are then separated by electrophoresis the position of the TBP is revealed by a peak of radioactivity between the α_1 and α_2 proteins when the radioactivity is scanned along the paper Albright and his colleagues have shown that the capacity of TBP to bind thyroxine is limited If the amount of thryoxine in the serum is increased to a point beyond which the binding sites on TBP are saturated, thyroxine begins to attach itself to albumin Since thyroxine bound to TBP is exchangeable with free thyroxine added to the serum this results in a progressive displacement of radioactivity from TBP to albumin Larson and Albright (1955) have shown that tri iodothyronine is not so firmly bound to TBP as thyroxine is Therefore it does not displace radioactive thyroxine molecules from their binding sites so readily

All these effects can be demonstrated in rat s serum though they are shown more clearly if a borate buffer is used for the electro-phoresis instead of a barbiturate buffer Plate XXV, Figure 10 shows a sample of rat s serum separated by paper electro-phoresis in a borate buffer and stained with naphthalene black

The A_1 band (corresponding to the α_1 fraction in human serum) has merged with the rear of the albumin band and is well separated from the A_2 band (corresponding to the α_2 fraction in human serum) Curve (*a*) shows how the radioactivity is distributed along the paper when the serum is taken from rats in which the endogenous hormone is labelled with ^{131}I Curves (*b*) to (*d*) show how the radioactivity is displaced on to the albumin as increasing amounts of non radioactive thyroxine are injected into the rats before the blood samples are taken Curve (*e*) shows that a large dose of non radioactive tri iodothyronine has a comparatively small displacing effect In this experiment, 500 micrograms of tri iodothyronine caused less displacement than 100 micrograms of thyroxine (curve (*e*) compared with (*c*))

It is reasonable to suppose that one of the factors which determine the biliary clearance rate of thyroxine is the ease with which the liver can detach thyroxine from the serum proteins If this is so, the clearance rate would be higher when the thyroxine was attached to albumin by weak bonds than when it was attached to TBP by strong bonds, and any factor tending to increase the proportion of the circulating thyroxine bound by albumin would increase the observed clearance rate On this hypothesis, the rise in clearance rate with increasing doses of thyroxine is due to the displacement of thyroxine from TBP to albumin, as shown in Plate XXV, Fig 10 It is also consistent with this explanation that tri iodothyronine, which does not displace thyroxine on to albumin very readily, has such a small effect on the biliary clearance rate of endogenous thyroid hormone

This hypothesis requires much more experimental evidence, especially of a quantitative kind, before it can be accepted However, it provides a possible explanation for what Gross and Leblond called an overflow when they observed that the proportion of a dose of thyroxine excreted in the faeces was greater with a large dose than with a small one

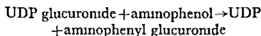
Significance of the Enterohepatic Circulation of Thyroid Hormone

Now I should like to consider the biliary excretion of thyroid hormone from a wider point of view and to see how far we can

go towards explaining it As physiologists we should, it seems to me, attempt two sorts of explanation—one which tells us how it works, and another which tells us what use it is to the body as a whole With this in mind, we have to consider two distinct and perhaps independent, processes the conversion of thyroid hormone to the glucuronide and the mechanism whereby the liver is able to pump thyroid hormone from the plasma into the bile against a high concentration gradient

Formation of the glucuronide

To take first the formation of the glucuronide in the liver thyroid hormone is not the only substance conjugated in this way Several of the steroid hormones are converted to their glucuronides in the liver, and it has recently been shown by Billing and Lathe (1956) that bilirubin is excreted in the bile as a glucuronide The recent work of Dutton and Storey (1953 1955) has shown that the liver contains a conjugate of uridine diphosphate (UDP) with glucuronic acid and that many substituted phenols act as acceptors for the glucuronic acid portion of this conjugate Aminophenol has been used extensively as an acceptor, and there is now good evidence that the reaction



UDP glucuronide + aminophenol → UDP + aminophenyl glucuronide

is catalysed by an enzyme in the liver Since this enzyme is not very specific towards its substrate, it is possible that the glucuronides of thyroxine the steroid hormones and bilirubin are all formed by the same reaction and with the same enzyme This is very likely to be the case with thyroxine, since thyroxine is itself a substituted phenol [1]

With regard to the function of this enzyme, it is usually assumed that the synthesis of glucuronides including that of thyroxine serves to promote the excretion of toxic substances by making them more soluble in water The conjugation of

[1] Isselbacher (1956) has recently shown that microsomes from guinea pig liver catalyse the formation of the glucuronides of tetrahydrocortisone testosterone thyroxine and nitrophenol when these substrates are incubated with UDP glucuronic acid This makes it almost certain that all these glucuronides are synthesized by the reaction described by Dutton and Storey

The A_1 band (corresponding to the α_1 fraction in human serum) has merged with the rear of the albumin band and is well separated from the A_2 band (corresponding to the α_2 fraction in human serum) Curve (*a*) shows how the radioactivity is distributed along the paper when the serum is taken from rats in which the endogenous hormone is labelled with ^{131}I Curves (*b*) to (*d*) show how the radioactivity is displaced on to the albumin as increasing amounts of non radioactive thyroxine are injected into the rats before the blood samples are taken Curve (*e*) shows that a large dose of non radioactive tri iodothyronine has a comparatively small displacing effect In this experiment, 500 micrograms of tri iodothyronine caused less displacement than 100 micrograms of thyroxine (curve (*e*) compared with (*c*))

It is reasonable to suppose that one of the factors which determine the biliary clearance rate of thyroxine is the ease with which the liver can detach thyroxine from the serum proteins If this is so, the clearance rate would be higher when the thyroxine was attached to albumin by weak bonds than when it was attached to TBP by strong bonds, and any factor tending to increase the proportion of the circulating thyroxine bound by albumin would increase the observed clearance rate On this hypothesis, the rise in clearance rate with increasing doses of thyroxine is due to the displacement of thyroxine from TBP to albumin, as shown in Plate XXV Fig 10 It is also consistent with this explanation that tri iodothyronine, which does not displace thyroxine on to albumin very readily, has such a small effect on the biliary clearance rate of endogenous thyroid hormone

This hypothesis requires much more experimental evidence, especially of a quantitative kind, before it can be accepted However it provides a possible explanation for what Gross and Leblond called an overflow when they observed that the proportion of a dose of thyroxine excreted in the faeces was greater with a large dose than with a small one

Significance of the Enterohepatic Circulation of Thyroid Hormone

Now I should like to consider the biliary excretion of thyroid hormone from a wider point of view and to see how far we can

intestine the effect of an increase in the biliary clearance rate must be to increase the rate of removal of thyroxine from the body by excretion *in the faeces* It is probable, therefore that the biliary route of excretion acts as a mechanism for the rapid adjustment of the concentration of thyroid hormone in the plasma The observations of Gross and Leblond (1950) show that this mechanism comes into play when a large dose of thyroxine is injected into a rat but it has yet to be proved that it also acts under natural conditions when the fluctuations in the plasma concentration of thyroxine must be far smaller than those induced experimentally (1 microgram of thyroxine is nearly twice the amount present in the whole of a rat's plasma)

The experiments on the binding of thyroxine by the serum proteins provide a possible explanation for the way in which the *biliary excretion of thyroxine is varied But here again on the* evidence available at present, it cannot be assumed that variations in the proportions of thyroxine bound to albumin and TBP also occur under physiological conditions

If the function of the biliary excretion of thyroid hormone is to adjust the amount of hormone in the circulation it is certainly not the only means of regulation available to the body The interaction between the thyroid and the pituitary tends to maintain a constant concentration of hormone in the plasma It has also been shown that the rate of breakdown of thyroid hormone increases when the plasma concentration is raised (Berson and Yalow 1954) This no doubt is the reason why biliary obstruction does not cause symptoms of an excess of thyroid hormone

SUMMARY

The liver extracts thyroxine from the plasma and excretes it in the bile partly as free thyroxine and partly as the glucuronide of thyroxine Some of the glucuronide is hydrolysed in the intestine and the resulting free thyroxine reabsorbed into the circulation The rest is lost in the faeces If the circulating thyroxine is labelled with ^{131}I the amounts of organic iodine taking part in each phase of the enterohepatic circulation can

exogenous phenols with glucuronic acid may well have evolved as a protective mechanism, but it is surely a contradiction in terms to speak of the detoxication of a hormone that the body synthesizes and that is required for normal health As I shall explain below, the formation of thyroxine glucuronide may help to regulate the concentration of thyroid hormone in the plasma Perhaps we should also keep in mind the possibility that it does not serve any special function, but is merely a consequence of the fact that the liver contains an enzyme that cannot distinguish thyroxine from alternative substrates

Active transport of thyroid hormone

At present, nothing is known about the mechanism by which the liver pumps thyroxine into the bile The active transport of several substances by other organs is known to depend on the formation of a reversible complex inside the cells Glucose, for instance is reversibly phosphorylated during its passage through the intestinal wall and the renal tubules By analogy with this, it might be thought that the formation of the glucuronide is an essential step in the transport of thyroxine across the liver some of the glucuronide leaking into the bile without being broken down However this seems to be very unlikely, because at high plasma concentrations of thyroxine when the bile/plasma concentration gradient is greatest the glucuronide may cease to appear in the bile

It is a curious fact that the liver excretes a great many substances with a molecular weight of a few hundreds (like thyroxine) and containing several iodine atoms Some of these like tetra iodophenolphthalein are so highly concentrated in the bile that they are used as contrast media for giving a radiographic outline of the biliary tract in humans It may be that these substances and the thyroid hormone are transported by the same mechanism If so, a comparison of their structures might suggest what this is

Finally, it remains to consider the function of the biliary excretion of thyroid hormone As we have seen, some of the thyroxine excreted in the bile is converted to the glucuronide Since thyroxine in this form is poorly absorbed from the

intestine, the effect of an increase in the biliary clearance rate must be to increase the rate of removal of thyroxine from the body by excretion in the faeces It is probable therefore that the biliary route of excretion acts as a mechanism for the rapid adjustment of the concentration of thyroid hormone in the plasma The observations of Gross and Leblond (1950) show that this mechanism comes into play when a large dose of thyroxine is injected into a rat, but it has yet to be proved that it also acts under natural conditions, when the fluctuations in the plasma concentration of thyroxine must be far smaller than those induced experimentally (1 microgram of thyroxine is nearly twice the amount present in the whole of a rat s plasma)

The experiments on the binding of thyroxine by the serum proteins provide a possible explanation for the way in which the biliary excretion of thyroxine is varied But here again, on the evidence available at present it cannot be assumed that variations in the proportions of thyroxine bound to albumin and TBP also occur under physiological conditions

If the function of the biliary excretion of thyroid hormone is to adjust the amount of hormone in the circulation it is certainly not the only means of regulation available to the body The interaction between the thyroid and the pituitary tends to maintain a constant concentration of hormone in the plasma It has also been shown that the rate of breakdown of thyroid hormone increases when the plasma concentration is raised (Berson and Yalow 1954) This no doubt is the reason why biliary obstruction does not cause symptoms of an excess of thyroid hormone

Summary

The liver extracts thyroxine from the plasma and excretes it in the bile, partly as free thyroxine and partly as the glucuronide of thyroxine Some of the glucuronide is hydrolysed in the intestine and the resulting free thyroxine reabsorbed into the circulation The rest is lost in the faeces If the circulating thyroxine is labelled with ^{131}I the amounts of organic iodine taking part in each phase of the enterohepatic circulation can

be measured under normal conditions In rats, about half the output of hormone from the thyroid is extracted by the liver and excreted in the bile Of this amount, nearly half is lost in the faeces Compared with the daily rate of synthesis of thyroid hormone, the enterohepatic circulation in humans seems to be much less extensive than it is in rats The biliary clearance rate of thyroxine increases as increasing amounts of thyroxine are injected into rats This may be because at high concentrations of thyroxine the thyroxine binding protein in the plasma is saturated, the excess of thyroxine becoming bound to albumin from which it is more easily removed by the liver The biliary excretion of thyroxine may help to regulate the concentration of hormone in the plasma

REFERENCES

ALBERT A and KEATING F R (1952) *Endocrinology* **51** 427
BERSON S A and YALOW R S (1954) *J clin Invest* **33** 1533
BILLING B H and LATHE G H (1956) *Biochem J* **63** 6P
DUTTON G J and STOREY I D E (1953) *Biochem J* **53** xxxvii
GORDON A H GROSS J O CONNOR D and PITT RIVERS R (195) *Nature Lond* **169** 19
GROSS J and LEBLOND C P (1950) *J biol Chem* **184** 489
ISSELBACHER K J (1956) *Recent Progr Hormone Res* **12** 134
JOLIOT F COURRIER R HOREAU A and SUE P (1944) *C R Acad Sci Paris* **218** 769
KENDALL E C (1919) *Endocrinology* **3** 156
KLITGAARD H M LIPNER H J BARKER S B and WINNICK T (1953) *Endocrinology* **52** 79
LARSON F C and ALBRIGHT E C (1955) *Endocrinology* **56** 737
LARSON F C DEISS W P and ALBRIGHT E C (1952) *Science* **115** 626
MYANT N B (1956a) *Clin Sci* **15** 227
MYANT N B (1956b) *Clin Sci* **15** 551
ROCHE J MICHEL R MICHEL O and TATA J (1954) *Biochim biophys acta* **13** 471
ROCHE J MICHEL R and TATA J (1954) *Biochim biophys acta* **15** 500
STOREY I D E and DUTTON G J (1955) *Biochem J* **59** 279
TAUROG A BRIGGS F N and CHAIKOFF I L (1951) *J biol Chem* **191** 29
TAUROG A BRIGGS F N and CHAIKOFF I L (1952) *J biol Chem* **194** 655

XIX

Resistance of Staphylococci to Antibiotics

MARY BARBER

JUST over fifteen years ago six patients with severe staphylococcal or streptococcal infections were treated with penicillin at the Radcliffe Infirmary, Oxford The results of this first clinical trial (Abraham *et al*, 1941) were so encouraging that with the help of the United States large scale production of penicillin was undertaken For a time the antibiotic was in very short supply, but those who were fortunate enough to obtain samples acclaimed it with enthusiasm In particular staphylococcal infection appeared to have been brought under control Intensive research for fresh antibiotics with a wider range of activity followed and when in 1944 the newly discovered antibiotic streptomycin was found to show promise in the treatment of tuberculosis (Schatz Bugie and Waksman 1944), it appeared possible that infectious diseases would be wiped out and bacteriology cease to be part of the medical curriculum

The more percipient remembered, perhaps, the warnings of Kossiakoff (1887) and Ehrlich (1908) that bacteria are infinitely adaptable to antiseptics But the era of antibiotics had arrived the great resources of the American drug industry were behind their production newspaper reports gave them world wide publicity and patients insisted on being treated with the latest one In Great Britain some control was kept and at least a medical prescription was necessary but in the U S A manufacture and sale ran riot so that penicillin was even incorporated into toothpaste and chewing gum

Today we reap the results of our haphazard use of these powerful drugs in the ever increasing incidence of infective processes resistant to antibiotic treatment Fortunately, however, although nearly all species of bacteria appear capable, under certain conditions, of yielding strains with an increased resistance to all the antibiotics, in clinical practice this has only become a major problem in the case of infections due to *Staph pyogenes* or *Mycobact tuberculosis* The former organism, the subject of this lecture, has shown such hardihood and adaptability in the face of attack that in most general hospitals today staphylococcal infection is the major bacteriological problem

LABORATORY ASPECTS

TYPES OF RESISTANCE

It might well be said of the staphylococcus my name is legion, for we are many' Drug resistant variants of this species differ, both in the way in which they respond to the antibiotic and in the presence or absence of associated changes These differences are not only determined by the antibiotic used and the environmental conditions, but, as Barber (1953b) has shown, a single strain of *Staph pyogenes* serially passaged in the presence of a single antibiotic may yield several types of drug resistant variant in a single experiment Indeed, this can be seen at a glance if a penicillin gradient plate is flooded with a broth culture of a *penicillin sensitive staphylococcus (cf* Plate XXVI, Figure *1*)[1]

With regard to their response to the antibiotic, three main types of resistance are encountered (1) *drug tolerance*, (2) *drug dependence*, (3) *drug destruction* In the first case the variant is capable of growing in the presence of an increased concentration of the antibiotic, although the latter is unchanged and retains full activity for other bacteria *in the second the strain* will not grow in the absence of the antibiotic, in the third the variant produces an enzyme (e g penicillinase), or some other type of antagonist, capable of destroying the antibiotic

Of the associated changes occurring in drug resistant variants the commonest is a reduction in growth rate, usually

[1] The plates referred to in this lecture will be found between pp 336–7 and 352–3

resulting from an increase in the lag phase This may lead to an apparent increase in sensitivity to other antibiotics (Chandler, Davidson, Long and Monnier, 1951, Monnier and Schoenbach, 1951, Barber, 1953a) Many of these slow growing variants show colonial changes, particularly in relation to size opacity and pigmentation (Plate XXIX, Figures 7 8) (cf Barber, 1953b and c, Barber and Burston, 1955) Morphological changes are also quite common (Rake *et al*, 1944 Blair *et al* 1946, Barber, 1953b) and what is of great importance clinically, most of the strains showing a decreased rate of growth and an abnormal morphology are of lowered virulence Finally a basic alteration in metabolism may be the fundamental reason for resistance

The rate of emergence of drug resistant staphylococci varies very much with the different antibiotics With streptomycin the emergence is so rapid that they frequently appear within two to three days of the onset of treatment of an infective process due to an initially streptomycin sensitive staphylococcus The rate is only slightly less with erythromycin and novobiocin On the other hand, with penicillin chloramphenicol, the tetra cyclines and vancomycin, drug resistant staphylococci occur relatively slowly and are unlikely to appear during the treat ment of a single clinical case

PENICILLIN

Penicillin resistant staphylococci (Plates XXVI, XXVII, Figures 2–4) fall clearly into the two groups, penicillin tolerant or peni cillin destructive Penicillin dependent strains have been des cribed (Barber 1953b), but are so unstable that their isolation is comparatively rare *Penicillin tolerant strains* are readily isolated by serial passage in the presence of penicillin *in vitro*, but are un stable in the absence of the antibiotic They are rarely en countered in clinical practice, probably because they grow more slowly than typical staphylococci and are of reduced virulence (Rake *et al* 1944 Spink Ferris and Vivino 1944, Blair *et al* 1946) Gale and his colleagues (Gale and Taylor, 1947 Gale and Rodwell 1949) have shown that staphylococcal variants of this type are capable of synthezising their amino acid requirements from ammonia and a carbon source whereas

Today we reap the results of our haphazard use of these powerful drugs in the ever increasing incidence of infective processes resistant to antibiotic treatment Fortunately, how ever, although nearly all species of bacteria appear capable, under certain conditions, of yielding strains with an increased resistance to all the antibiotics, in clinical practice this has only become a major problem in the case of infections due to *Staph pyogenes* or *Mycobact tuberculosis* The former organism, the sub ject of this lecture, has shown such hardihood and adaptability in the face of attack that in most general hospitals today staphylococcal infection is the major bacteriological problem

LABORATORY ASPECTS

Types of Resistance

It might well be said of the staphylococcus 'my name is legion for we are many Drug resistant variants of this species differ, both in the way in which they respond to the antibiotic and in the presence or absence of associated changes These differences are not only determined by the antibiotic used and the environ mental conditions, but, as Barber (1953b) has shown, a single strain of *Staph pyogenes* serially passaged in the presence of a single antibiotic may yield several types of drug resistant vari ant in a single experiment Indeed, this can be seen at a glance if a penicillin gradient plate is flooded with a broth culture of a penicillin sensitive staphylococcus (cf Plate XXVI, Figure 1) [1]

With regard to their response to the antibiotic, three main types of resistance are encountered (1) *drug tolerance*, (2) *drug dependence*, (3) *drug destruction* In the first case the variant is capable of growing in the presence of an increased concentra tion of the antibiotic although the latter is unchanged and retains full activity for other bacteria in the second the strain will not grow in the absence of the antibiotic, in the third the variant produces an enzyme (e g penicillinase) or some other type of antagonist capable of destroying the antibiotic

Of the associated changes occurring in drug resistant vari ants the commonest is a reduction in growth rate, usually

[1] The plates referred to in this lecture will be found between pp 336–7 and 352–3

Plate XXVI

Penicillin resistant strains *Staph pyogenes*

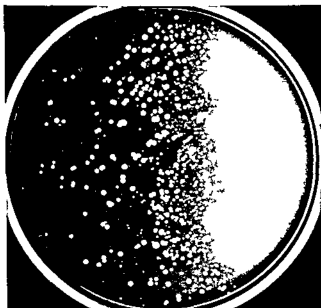
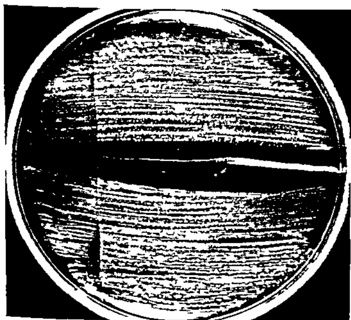


Fig 1 Graduate plate (0 1 u/ml penicillin in top layer) flooded with penicillin sensitive strain Note variety of penicillin resistant colonies



cillin tolerant strain plated out on penicillin ditch plate (10 ditch) Central streak is a penicillin sensitive strain

(From Barber 1953c)

penicillin sensitive strains have to assimilate them from the surrounding medium, a process which is blocked by penicillin

Penicillin destroying strains, on the other hand, are almost invariably the cause of penicillin resistant staphylococcal infection, but resistant strains of this type do not occur when penicillin sensitive strains of staphylococci are subcultured in penicillin *in vitro* Penicillin destroying strains of *Staph pyogenes* are fully virulent and indeed resemble penicillin sensitive strains in all respects except the capacity to produce penicillinase (Barber, 1947a) This capacity is a relatively permanent characteristic, but there is a tendency for penicillin destroying strains to yield a proportion of penicillin sensitive cells, which completely fail to produce the enzyme (Barber, 1949, Bondi *et al* 1953 Fairbrother *et al*, 1954)

It is important to remember that the degree of resistance of penicillin destroying staphylococci depends on the number of bacteria present, and in laboratory tests may vary many hundredfold, according to the size of inoculum used (Luria 1946, Barber, 1947a) If such a culture be plated out on a penicillin ditch plate it will be seen that the area of inhibition increases as the number of bacteria decreases (Plate XXVI, Figure 2) This makes it impossible for laboratory reports to state the precise concentration of penicillin by which such an organism may be inhibited Gilson and Parker (1948) suggested a double sensitivity test in broth using two different sized inocula Barber and Whitehead (1949), using such a test found that when the inoculum was approximately a million organisms all of 17 strains grew in the presence of 12 5 u/ml penicillin or more, whereas when the inoculum was only a thousand bacteria the greatest concentration of penicillin permitting growth ranged from 4 to 0 03 u/ml Since however, in most cases of staphylococcal infection the organisms are present in large numbers, it is probable that a single test using a large inoculum is sufficient for routine purposes

Streptomycin

Nearly all species of bacteria are capable of yielding variants showing gross resistance to streptomycin after only a few

PLATE XXVI

Penicillin resistant strains *Staph pyogenes*



FIG 1 Graduate plate (0 1 u/ml penicillin in top layer) flooded with penicillin sensitive strain Note variety of penicillin resistant colonies

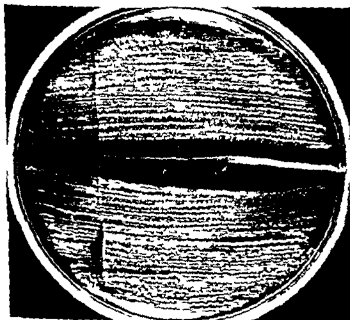


FIG 2 Penicillin tolerant strain plated out on penicillin ditch plate (10 u ml pen in ditch) Central streak is a penicillin sensitive strain

(From Barber 1953c)

Plate XXVII

Penicillin resistant strains *Staph pyogenes*

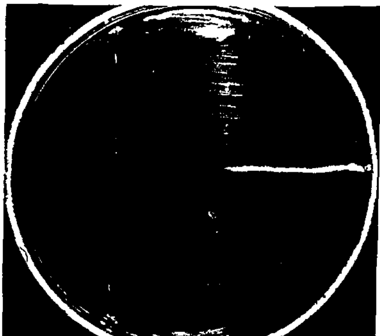


Fig 3 Penicillin dependent strain plated out on penicillin ditch plate (10 u/ml pen in ditch) Central streak is a penicillin sensitive strain (From Barber 1953c)

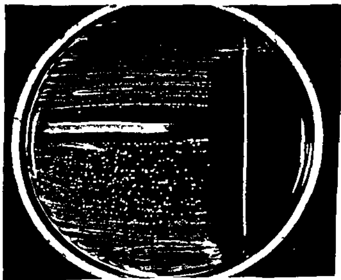


Fig 4 Penicillinase producing strain plated out on penicillin ditch plate (10 u/ml pen in ditch) Central streak is a penicillin ...ve strain ...er 1953c)

subcultures in the presence of the antibiotic, and staphylococci are no exception Streptomycin dependent strains are also frequently encountered (Plate XXVIII, Figures 5, 6) The resistance is not associated with destruction of the antibiotic and appears to be of the same type in *in vitro* experiments and clinical practice Streptomycin resistant staphylococci sometimes show changes in growth rate and virulence, but many, at any rate, resemble penicillin sensitive strains in these respects Umbreit and his colleagues have shown that oxalacetate pyruvate condensation is inhibited by streptomycin, and that streptomycin resistance is usually associated with a change in the pathway of pyruvic acid oxidation (Oginsky *et al*, 1949, Smith *et al*, 1949)

Chloramphenicol

Chloramphenicol resistant staphylococci are yielded comparatively slowly As in the case of stretomycin resistant strains, resistance is not associated with destruction of the antibiotic, although a chloramphenicol destroying enzyme has been described in some species of bacteria (Smith and Worrel, 1949 and 1950) Chloramphenicol resistant strains of *Staph pyogenes* of apparently full virulence have been isolated from infective processes, but are relatively uncommon now that the use of this antibiotic is limited, on account of its toxicity

Tetracyclines

Tetracycline resistant staphylococci, like those resistant to chloramphenicol, appear slowly (Paine Collins, Finland 1948 Finland Collins, Paine, 1948), but owing to the extensive use of the tetracycline antibiotics they are now quite common in hospitals The resistance is not associated with antibiotic destruction, but resistant strains appear to be fully virulent There is almost complete cross resistance between chlortetracycline oxytetracycline and tetracycline, although chlortetracycline is slightly more effective against tetracycline resistant variants as it is against sensitive strains (Love *et al* 1954)

Erythromycin

Staphylococci yield strains resistant to erythromycin more readily than to any other antibiotic except streptomycin

(Hobson, 1954), and the spread of such strains in a hospital where the use of erythromycin is extensive may be very rapid (Lepper *et al*, 1954) Erythromycin resistant strains do not destroy the antibiotic but they may be fully virulent, as shown by the work of Wise and his colleagues of Minneapolis (Wise *et al*, 1955) who record twenty three patient with erythromycin resistant staphylococcal infection, six of whom died The strains show almost complete cross resistance with magnamycin, and to some extent with the two newer antibiotics, spiramycin, isolated in the laboratories of Rhone Poulenc (Chabbert, 1955), and E 129, isolated in the Glaxo Laboratories (cf Garrod and Waterworth 1956)

Recently Isolated Antibiotics

Early reports on novobiocin (albamycin, cathomycin, carbomycin, streptonivicin), independently isolated in 1955 in three different laboratories of the United States, show that strains of staphylococci resistant to this antibiotic occur fairly readily both *in vitro* and in clinical practice (Nichols and Finland, 1956) With another recently discovered antibiotic, vancomycin, resistant strains are much less easily obtained (Garrod and Waterworth 1956)

Mode of Origin

A great deal of literature has been devoted to a discussion of the mode of origin of drug resistant variants of bacteria and geneticists of different schools of thought have tended to take rigidly opposing views as to the effect of the environment Some authors appear to believe that all drug resistant variants arise by a process of *spontaneous mutation in which the presence* of the antibiotic has only a selective action whereas others claim that the change occurs by a process of adaptation to the environment Thus spontaneous mutation and adaptive change have become sharply contrasted theories, and in the words of Dean and Hinshelwood (1953) doctrinal *arguments of a distinctly* scholastic flavour have tended to create antitheses which in fact are unreal

To the unindoctrinated it seems clear that mutation and adaptation are not mutally exclusive processes, and that

attempts to find a single explanation for the mode of origin of all the many types of drug resistant bacteria are futile That bacteria do yield what appear to be spontaneously occurring mutants is indisputable, they are also very labile organisms, which are peculiarly susceptible to changes in the environment

In a few laboratory experiments bacterial variation has been induced by the specific transfer of properties, including antibiotic resistance, from one cell to another Examples of this are the transformation of pneumococcal types by the nucleic acid fraction of specific nucleoproteins, the sexual fusion and recombination of certain strains of *Bact coli* and transduction of properties by bacteriophage in Salmonella species A detailed description of these reactions is beyond the scope of this lecture but antibiotic resistance may be specifically induced by these processes and those interested are referred to the review by Austrian (1952)

Antibiotic tolerant Staphylococci In most cases antibiotic resistant variants arise by a discontinuous process in which the population is not heterogeneous, suggesting spontaneous mutation Thus in the first instance a few variant cells appear, which if not favoured by the environment are likely to be overgrown The presence of the antibiotic however, has a selective action so that the few antibiotic resistant mutants are favoured at the expense of the rest of the population until a pure culture of antibiotic resistant cells is obtained

Since naturally occurring mutants have a random distribution their number will vary in experiments with different cultures of the same organisms but should be similar in different samples of the same culture (cf Luria and Delbruck 1943) Demerec (1945 and 1948) found this to be the case with streptomycin resistant and penicillin resistant variants of *Staph pyogenes* Further confirmation of the spontaneous origin of antibiotic resistant variants has come from the work of Lederberg and Lederberg (1952) who by means of a replica plating method were able to demonstrate the pre existence of mutants before their selection by antibiotic in the environment

Penicillin destroying Staphylococci Although penicillinase production by staphylococci is a fairly stable characteristic peni

cillin resistant strains of this type, when subcultured in the laboratory, tend to yield an increasing proportion of cells which have completely lost the capacity to produce the enzyme (Barber, 1949, Fairbrother, Parker and Eaton, 1954) The change appears to be a sudden complete loss and suggests spontaneous mutation It is quite possible, therefore, that cells which have gained the capacity to produce penicillinase arise by a process of spontaneous mutation from penicillin sensitive staphylococci Since, however, this variation has not been un equivocally demonstrated either *in vitro* or *in vivo*, the mutation must be a rare one

Recently Barber (1957) has shown that certain penicillin sensitive strains of *Staph pyogenes* yield variants with very weak penicillin destroying activity (less than one hundredth that of a typical clinical strain) when passaged over very long periods in the presence of very low concentrations of pencillin Whether this weak activity is due to a penicillinase similar to that of the strains so common in clinical practice, and can be increased by further passage remains to be determined but, if so, sub bacteriostatic concentrations of penicillin in the tissues or abscesses of penicillin treated cases may be an important factor in the emergence of penicillin resistant staphylococcal infection Moreover, transfer of such strains from case to case in hospitals may well constitute a method of passaging strains in sub bacteriostatic concentrations of penicillin

EPIDEMIOLOGY OF ANTIBIOTIC RESISTANT STAPHYLOCOCCAL INFECTION

Whatever the ultimate mode of origin of antibiotic resistant bacteria, it is clear that the increased incidence of antibiotic resistant staphylococcal infection in hospitals is due to the selection of drug resistant strains by the use of antibiotics, and the spread of these strains among patients and nurses in the hospital environment Indeed it is no exaggeration to say that hospitals where the use of antibiotics is widespread have become breeding grounds for a few strains of antibiotic resistant *Staph pyogenes*, the spread of which is not prevented by what is usually regarded as satisfactory aseptic technique

EARLY REPORTS OF ANTIBIOTIC RESISTANT STAPHYLOCOCCAL INFECTION

Barber (1947b) and Barber and Rozwadowska Dowzenko (1948) first drew attention to the rapid and progressive increase in the incidence of penicillin resistant staphylococcal infection in a large London hospital between the years 1946 and 1948 They found that the percentage of cases of staphylococcal in fection in this hospital which were resistant to penicillin was 14 per cent in 1946, 38 per cent in 1947 and 59 per cent in 1948 Similar reports soon followed from other hospitals in England (Forbes, 1949, Barber and Whitehead 1949, Barber *et al*, 1949 Cairns and Summers, 1950) Australia (Rountree and Thomson 1949) the United States (Nichols and Need ham, 1949, Spink, 1951, Kirby and Ahern, 1953) France (Chabbert *et al*, 1953), Scandinavia (Eriksen 1952 Vogelsang, 1953 Wallmark 1954a), the U S S R (Glasman 1953) Chile (Valenzuela and Vaccaro, 1954), Egypt (El Ghoroury 1954) and Finland (Anttonen, 1955) Today penicillin resis tant infection is quite common among hospital out patients (Wallmark 1954b, Rees *et al*, 1955 Fairbrother 1956) and even in general practice (Roodyn, 1954)

The story with minor variations has been repeated with the newer antibiotics wherever their use has been widespread, and today staphylococcal infection in hospitals is frequently due to strains resistant to penicillin, streptomycin and all the tetra cycline compounds (Rountree and Thomson, 1952, Lowbury *et al*, 1952 Clarke *et al* 1952, Kirby and Ahern, 1953, Barber and Burston 1955, Brodie *et al* 1955 Finland 1955) It is interesting to note however that since the use of chloramphenicol has been reduced because of possible bone marrow damage strains resistant to this antibiotic have become comparatively rare (cf Kirby and Ahern 1953 Finland, 1955 Barber and Burston, 1955)

At present the incidence of erythromycin resistant staphy lococci is negligible in most hospitals in this country It is clear however that this will not remain true unless the use of erythromycin be strictly controlled As indicated earlier

staphylococci readily yield strains resistant to it and this is not simply a laboratory phenomenon Thus Thomson *et al* (1956) report that in 6 of 71 cases of staphylococcal infection treated with erythromycin, resistance developed during treatment

GENERAL POPULATION OUTSIDE HOSPITALS

Until recently antibiotic resistant staphylococcal infection was regarded as almost exclusively a hospital problem Unfortunately *this is no longer the case with regard to penicillin* resistant infection, although the incidence is still much less frequent outside hospitals, and infection resistant to other antibiotics is still largely confined to these institutions The first change was noted among hospital out patients Early workers reported that the incidence of antibiotic resistance infection among this group was small (Forbes, 1949, Barber and Whitehead, 1949) In 1955, however, Rees *et al* found that 21 5 per cent of 200 strains of *Staph pyogenes* isolated from acute infections attending a casualty department were penicillin resistant, and Fairbrother (1956) found an even higher incidence among out patients

The problem has now reached the general practitioner Roodyn (1954), working at a health centre, found that 23 per cent of 93 strains of staphylococci isolated from boils and styes were penicillin resistant and in a later study (Roodyn, 1956) he found that 30 9 per cent of strains isolated from the anterior nares of patients suffering from recurrent infection were resistant to penicillin

Finally Rountree and her colleagues (Rountree and Rheuben, 1956) in Sydney have demonstrated the increasing incidence of penicillin resistant strains of *Staph pyogenes* isolated from nasal swabs of blood donors Thus the percentage of donors carrying penicillin resistant strains of *Staph pyogenes* was 3 5 in 1951, 6 5 in 1954 and 13 in 1955 These figures speak for *themselves and illustrate what will happen with the newer* antibiotics if their use is not controlled The same investigator (Rountree, 1956) studied the nasal flora of natives of the Wabag region of New Guinea a group of people which has had very little contact with so called civilized communities She found

that 23 of 120 people carried *Staph pyogenes* in the nose, but none of these strains was antibiotic resistant, although 18 of 100 coagulase negative strains isolated were penicillinase producers

INFECTION IN MATERNITY DEPARTMENTS

In 1949 Barber, Hayhoe and Whitehead studied a maternity hospital where staphylococcal infection had become a serious problem *Staph pyogenes* was isolated from 42 infective processes in infants, in 39 instances penicillin resistant strains were found and all but two of the typable resistant strains were phage type 52A Throughout a period of eight months 56–58 per cent of the nursing staff of this hospital were found to be nasal carriers of *Staph pyogenes*, 70 to 80 per cent of these carriers carried penicillin resistant strains and again nearly all the typable resistant strains were type 52A This picture remained constant in spite of the fact that the nurses themselves were continuously changing as the hospital was a training school for pupil midwives who came for a period of one to three months When the nasal carriage rate of penicillin resistant *Staph pyogenes* was studied in relation to the length of time the nurses had been in the hospital it was found that the two went closely together Thus only 24 per cent of nurses who had been in the hospital for a fortnight or less carried such strains whereas the incidence among nurses who had been there three months or more was 67 per cent

It seemed clear that a single strain of *Staph pyogenes* of phage type 52A and resistant to penicillin was spreading round this hospital and was responsible for the sepsis

In a later study (Barber *et al* 1953) of a maternity unit in which sepsis was uncommon and relatively trivial it was found, nevertheless that 61 per cent of the nursing staff were nasal carriers of penicillin resistant *Staph pyogenes* and the majority of typable strains were phage type 52A Moreover 65 per cent of the infants left the hospital carrying penicillin resistant strains of *Staph pyogenes* in the anterior nares which were similar in phage type to those isolated from the nurses, and in the majority of cases the nasal flora of the infants differed from that of their

respective mothers No strains resistant to antibiotics other than penicillin were encountered

Laboratory tests indicated that nearly all the carrier strains isolated from nurses and babies were active in the production of coagulase and α haemolysin It must, therefore, be assumed that potentially virulent strains of penicillin resistant *Staph pyogenes* were spreading round the department Why then there was no severe sepsis remains a problem and will be discussed in the section devoted to the virulence of staphylococci and the resistance of the host Further studies of the same maternity unit two years later (Barber and Burston, 1955) yielded an essentially similar picture, and even at this date strains resistant to antibiotics other than penicillin were extremely rare

It is impossible to summarize in one lecture the work of all investigators in this field but mention of a few studies will bring out the salient features Colbeck (1949) described an outbreak of breast abscesses in the maternity units of four general hospitals in Winnipeg caused by a single penicillin resistant strain of *Staph pyogenes* typed by a specific phage isolated early in the investigation He isolated a similar staphylococcus from the nose or throat of 26 2 per cent of the infants in one of the four hospitals, and suggested that this was the reservoir of infection

An interesting series of papers on staphylococcal infection in the newborn have come from a group of workers in Edinburgh Forfar *et al* (1953) reported on the sepsis occurring in two large maternity units in 1951 Just over 6 per cent of the babies in each of these hospitals developed infections of skin or conjunctiva most of which were staphylococcal and a number of cases developed more serious sepsis, often after returning home Of 56 strains of *Staph pyogenes* isolated from these cases 76 7 per cent were resistant to penicillin 12 5 per cent to streptomycin 17 4 per cent to chloramphenicol and 4 5 per cent to chlortetracycline Since no strains resistant to either of the two latter antibiotics were isolated from members of the nursing staff, it was assumed that cross infection was a more important source of spread than infection from nursing staff

In 1955 the same group of workers (Edmunds *et al* 1955) recorded the characteristics of pathogenic staphylococci isolated

in the environment of newborn infants born in hospital and at home Three different maternity units were studied in all of which there was a high incidence of mild sepsis but no serious infection Nevertheless, as in previously described studies (Barber *et al*, 1953, Barber and Burston, 1955) spread of *Staph pyogenes* to infants was similarly frequent in all three hospitals The staphylococcal advantages of being born at home were clearly demonstrated and this despite the fact that there was a high nasal carriage rate of *Staph pyogenes* among the nurses attending

Finally mention must be made of the epidemic of staphylococcal pneumonia in the newborn described by Beavan and Burry (1956) from Christchurch, New Zealand The outbreak occurred in the maternity annex of a large private hospital where the babies were nursed in a single nursery designed to accommodate thirty cots, and sometimes allowed to contain a larger number At the onset of the epidemic minor staphylococcal sepsis was considerably more prevalent than in other maternity units in the city The epidemic included eight fatal cases and autopsy showed a haemorrhagic pneumonia, apparently resulting from inhalation with a tendency to suppuration and consistent and early involvement of the pleural cavity The staphylococci isolated were all resistant to penicillin and highly sensitive to erythromycin, but there was some variation in the degree of sensitivity to streptomycin, the tetracyclines and chloramphenicol Phage typing of a limited number of strains indicated that many if not all, were of the same phage pattern and belonged to Group III The authors point out the important lesson that 'minor staphylococcal infection in the nursery must not be taken lightly

Post operative Infection

Clarke *et al* (1952) were among the first to report the appearance and dissemination in a hospital of a staphylococcus resistant to penicillin *streptomycin chloramphenicol and the* tetracyclines As many as 20 of 25 infections from which this organism was isolated were post operative cases and clearly the result of hospital cross infection In a later paper the same

group of workers (Clarke *et al*, 1954) discuss the high incidence of cross infection with penicillin resistant *Staph pyogenes* in two surgical wards They demonstrated that dust suppression, by oiling floors and bedclothes, did not affect the incidence and concluded that contact spread is more important than are air borne bacteria

Barber and Burston (1955) analysed 100 cases of staphylococcal infection occurring in a general hospital between November 1954 and April 1955 Only 40 per cent of the strains isolated were sensitive to all antibiotics and nearly all of these were isolated from out patients or patients admitted with closed abscesses or recent acute infections There were 28 strains resistant to penicillin only 10 of these were from out patients and the rest came from all departments of the hospital Analysis of 24 strains resistant to penicillin, streptomycin and the tetracyclines, however, showed that 17 were from post operative infections and all the remaining 7 from patients admitted to hospital with chronic open infections

All the multiple resistant strains responsible for the post operative infection were phage group III and 12 were identical in phage type Nevertheless no single source of infection could be found, since the patients were in ten different wards and had been operated on by nine different surgeons in three operating theatres As in most large hospitals however, clean and septic cases were admitted to the same wards and the seven patients with open chronic infection due to antibiotic resistant strains of similar phage type were in five of the ten wards concerned Cross infection from these and similar cases seemed therefore, a probable explanation

An important study of cross infection with penicillin resistant staphylococci in a thoracic surgery unit has recently come from Middlesbrough (Blower *et al* 1955) These workers report that in 1952 the incidence of infection was 10 9 per cent of all cases and as a result the unit was temporarily closed The staphylococci isolated were resistant to penicillin only and were of various phage types all in Group III They attributed this infection to a number of causes notable among which were lack of facilities to isolate infected patients, and in the theatre

the use of inefficient ventilation, unsterilized blankets and excessive activity of the staff

Super infection during Antibiotic Treatment

In any patient treated with one of the broad spectrum antibiotics the normal bacterial flora of the body is practically eliminated In such patients many parts of the body, particularly the intestinal tract, are unusually prone to attack by antibiotic resistant staphylococci in the environment This may result in infections of unusual type and severity Finland and his colleagues in Boston, Massachusetts, were the first to draw attention to this In 1951 (Jackson *et al*) they studied 91 cases of pneumonia, mostly pneumococcal, treated with oxytetracycline and found that in 48 *Staph aureus* replaced other bacteria and became the predominant organism in the sputum during the course of treatment In the whole series there were only seven deaths and staphylococcal super infection accounted for four Such unusual conditions as staphylococcal dysentery' (Womack *et al*, 1952) and staphylococcal scarlet fever (Hazen *et al*, 1951), developing during antibiotic therapy, were reported from the same hospital (see also Finland 1951)

Hay and McKenzie (1954) reported the side effects of oxytetracycline therapy in 603 cases (mostly children) of gastro intestinal infection in Glasgow Twenty one developed staphylococcal infections consisting of 2 cases of fatal gastro entero colitis 7 cases of scarlet fever, 8 cases of sore throat without a rash, 1 case of balanitis and scarlatina and 3 urinary infections Bacteriological investigations were not done in all cases but coagulase positive staphylococci were isolated from 14 Of 12 strains tested all were resistant to penicillin and the tetracyclines and 2 also to streptomycin and 2 to chloramphenicol 5 strains were phage typed and though all were Group III there were 3 different types Thus infection was not due to spread of a single strain but the danger of broad spectrum antibiotics is clear, and as the authors point out they should not be used in minor illnesses

A similar report comes from Brodie and his colleagues in St Andrews (Brodie *et al*, 1955) They studied the complications

occurring in 378 children treated with tetracycline for bacillary dysentery Antibiotic resistant staphylococcal infection occurred in 48 and consisted of scarlet fever, sore throat without rash, enteritis or a combination These infections were apparently due to a 'hospital staphylococcus resistant to penicillin, streptomycin and the tetracyclines, and of a defined serological type

Pseudomembranous enterocolitis This fulminating cholera like condition was first observed by Finney in 1893 in a patient who had been operated on for pyloric obstruction Many reports of similar cases followed, and although in most instances the condition followed abdominal surgery, particularly operations of the colon (Pettet *et al*, 1954), a few cases occurred without any operation (Kleckner *et al*, 1952) The aetiology remained obscure, although many explanations were put forward Thus Reidel (1902) suggested that long operations with severe blood loss incapacitated the intestines, and Penner and Bernheim (1939) thought that post operative shock caused dilation of submucosal capillaries and venules, leading to oedema, haemorrhage and finally necrosis of the overlying mucosa

Since the widespread use of the broad spectrum antibiotics, sudden, often fatal, enterocolitis following surgical operations has been reported with increasing frequency (Jackson *et al*, 1951 Janbon *et al*, 1952 Terplan, 1953, Dearing and Heilman, 1953, Gardner 1953, Hay and McKenzie, 1954, Williams, 1954, Frame and Short, 1955, Fowler 1955) In most cases antibiotic resistant staphylococci have been incriminated and it is suggested that pre operative sterilization of the gut with antibiotics has left a free field for implantation with hospital strains of this organism In many cases examination of a Gram stained film of faeces is sufficient to establish the diagnosis, since the normal faecal flora is completely replaced by clumps of Gram positive cocci

Although some cases of post-operative enterocolitis are probably due to other bacteria, particularly *Cl welchii* (cf Howie *et al* 1953) or not associated with any pathogenic micro organisms (Williams and Pullan 1953), antibiotic resistant staphylococci appear to be the main aetiological agent in an

increasing number Recently Surgalla and Dick (1955) found that 30 of 32 strains of staphylococci isolated from the faeces of patients during a course of antibiotic produced an enterotoxin pathogenic for monkeys These authors suggest that the severity of post antibiotic enterocolitis, as opposed to staphylococcal food poisoning is due to the fact that in the absence of the normal gut flora staphylococcal enterotoxin production persists within the body

Staphylococcal Pneumonia in Infants Many workers have shown that the incidence of staphylococcal pneumonia in children has increased greatly since the introduction of the broad spectrum antibiotics Thus Prissick (1953) reported that in a children s hospital in Montreal the percentage of deaths due to this condition was 11 9 in 1951, whereas the highest figure in any previous year between 1942 and 1950 was 3 2 Wallman *et al* (1955) saw 55 cases (10 fatal) of staphylococcal pneumonia in infants at a children s hospital in Perth during a 29 month period in 1953 to 1955, whereas previously they regarded it as a rare disease Disney *et al* (1956), reporting a series of 35 infants with staphylococcal pneumonia in a hospital in Birmingham in 1953 and 1954, concluded that if their findings were representative, any infant under two years old with clinical signs of pneumonia or fluid in the chest has more than three chances to one of being infected with *Staph aureus* Most of the strains isolated in all these investigations were resistant to penicillin and many to other antibiotics Wallman *et al* (1955) commented on the relative ineffectiveness of the broad spectrum antibiotics in controlling the infection

Virulence and Bacteriophage Types

Staphylococcal infection appears to have increased both in frequency and severity since the introduction of antibiotics The gradual elimination of antibiotic sensitive strains and the selection of antibiotic resistant staphylococci mainly of a single phage group, together with cross infection from case to case, has, no doubt, led to the selection of increasingly virulent organisms Unfortunately it is difficult to demonstrate differences in the virulence of pathogenic staphylococci in the laboratory

so that evidence for increased virulence of a particular strain is circumstantial only A good deal of evidence of this nature is however, accumulating

Barber and Burston (1955) show an association between post operative and neonatal cases of infection with other infective processes rather than nasal carriers Both Rountree and Freeman (1955) and Beavan and Burry (1956) describe outbreaks of severe neonatal infection due to strains which had apparently become of increased virulence by passage among babies suffering from minor infections, which had been occurring with increasing frequency in the department for some time previously Brodie, Sommerville and Wilson (1956) found that newly arrived nurses did not become nasal carriers of *Staph pyogenes* while they were in the preliminary training school, although they were in close contact with more senior nurses who carried such organisms, but did so soon after entry into the wards, where, of course, they came into contact with infective processes

Many workers have recently claimed that different phage types of *Staph pyogenes* vary in virulence and in their capacity to cause certain types of infection Thus penicillin resistant strains of phage type 70 are the common cause of impetigo (Parker *et al*, 1955, Barrow, 1955) Alder *et al* (1955) found that most boils and nasal carriers yielded Group I strains, whereas wound infections were usually caused by Group III strains, although in the latter case the multiple antibiotic resistance of this group was probably a major factor Further examples are given by Anderson and Williams (1956) It seems probable, however, that all phage types of *Staph pyogenes* are of potentially high virulence once they become established in an institution

The association between certain bacteriophage types and antibiotic resistance is, however well established In 1949 Barber and Whitehead in England and Rountree and her colleagues in Australia found that penicillin resistant staphylococcal infection occurring in hospitals was due to strains of *Staph pyogenes* belonging to only one or two phage types (Barber and Whitehead, 1949 Rountree and Thomson, 1949), and

that many of the nursing staff carried similar strains in the anterior nares (Barber, Hayhoe and Whitehead, 1949, Rountree and Barbour, 1951) In the first cases described by both groups of workers, the strains belonged to phage group III, but in many later cases staphylococci of phage type 52A

TABLE 1 Bacteriophage Groups of Antibiotic sensitive and Resistant Strains of *Staph pyogenes*

Date	Antibiotic sensitivity	Bacteriophage Groups							
		% of strains isolated from medical and surgical wards				% of strains isolated from maternity units			
		I	II	III	N T	I	II	III	N T
1948	Resistant to Pen only	6	3	82	9	—	—	—	—
1949		10	0	70	20	60	0	0	40
1953		—	—	—	—	43	3	16	38
1955		34	17	48	0	69	8	19	4
1955	Resistant to Pen Strep	0	0	100	0	0	0	100	0
1955	Resistant to Pen Strep Tetra	0	0	100	0	0	0	100	0
1948–55	Antibiotic sensitive	31	21	19	29	32	25	24	19

Pen = Penicillin Strep = Streptomycin Tetra = Tetracycline antibiotics
N T = non typable

(Group I) were isolated, and penicillin resistant organisms of this type have since tended to predominate in many maternity units (cf Barber Hayhoe and Whitehead 1949 Barber *et al*, 1953 Barber and Burston, 1955) The proportion of penicillin resistant strains of phage group II is relatively small, but a moderate number are now being encountered (Williams *et al* 1953 Barber and Burston 1955)

Strains of *Staph pyogenes* of phage group III appear to yield drug resistant variants more readily than do strains of other phage groups and today most infective processes in British hospitals resistant to streptomycin or the tetracyclines are due to Group III strains The personal findings of the author with regard to the relationship of antibiotic sensitivity and phage group are summarized in Table 1 A similar association between phage group III and antibiotic resistance has been recorded

in hospitals in France (Fouaces and Lutz, 1953), Denmark (Wallmark, 1954a) and the U S A (Fusillo *et al*, 1954, Jackson *et al*, 1954)

RESISTANCE OF HOST

On the host side several factors need to be considered Of prime importance is the bacterial vacuum caused by treatment with broad spectrum antibiotics Since Pasteur posed the question—'la vie sans microbes est il possible?'—the role of the normal bacterial flora of the body has been studied extensively Today it is clear that whatever else these fellow travellers may do for us, they certainly act as a defence against invaders

In post operative cases tissue damage and irritation by sutures, starvation and shock may all play a part Dubos and his colleagues at the Rockefeller Institute have shown that a variety of procedures which disturb metabolism, particularly *starvation, increases the susceptibility of mice to infection* with staphylococci and other bacteria (Dubos *et al*, 1955 Dubos, 1955) In cases of staphylococcal pneumonia an antecedent virus infection may be a predisposing cause (Disney *et al*, 1956, Evans and Evans, 1956)

Finally mention must be made of specific antibodies Although most adults in this country have developed antibodies to staphylococci and many of its toxic products without becoming immune to infection, it seems probable that the administration of staphylococcal toxoid increases resistance to infection as well as the amount of circulating antitoxin (Dolman, 1933 1935, Parish, O Meara and Clark, 1934, Whitby 1936) Antibodies to coagulase (Lominski and Roberts, 1946), leucocidin (Valentine and Butler, 1939) and staphylococci themselves (Lyons 1937) have been demonstrated, but much further study is needed before their role in immunity can be assessed

PREVENTION

Control of Antibiotics The emergence of antibiotic resistant staphylococci in ever increasing numbers and distribution can only be prevented by controlling the use of antibiotics When the use of a particular antibiotic is stopped in a given institution the incidence of staphylococci resistant to it often falls quite

Plate XXVIII

Streptomycin resistant strains *Staph pyogenes* plated out on streptomycin ditch plates (100 µgm/ml streptomycin in ditch) Central streak is a streptomycin sensitive strain



Fig 5 Streptomycin sensitive strain (*upper half*) and streptomycin tolerant strain (*lower half*) derived from it after 6 passages in streptomycin

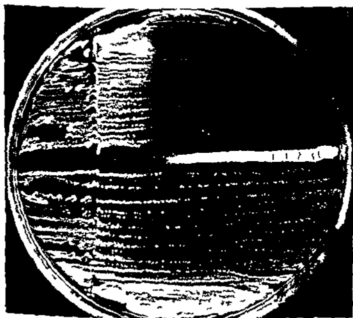


Fig 6 Streptomycin dependent strain (*upper half*) and strain favoured by streptomycin (*lower half*) (From Barber 1953c)

Plate XXIX

Antibiotic resistant strains *Staph pyogenes* showing associated changes

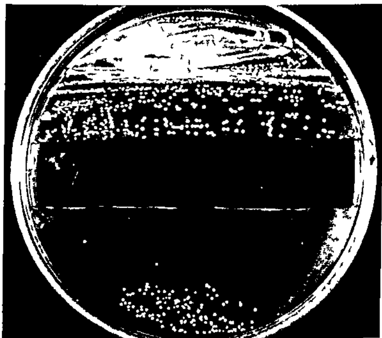


Fig 7 Antibiotic sensitive strain (*upper half*) and streptomycin resistant small colony variant (*lower half*)

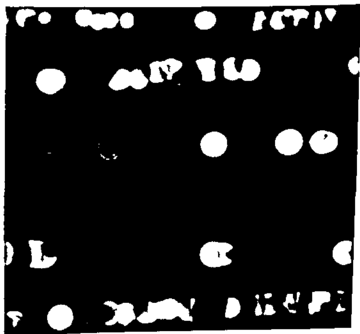


Fig 8 Strain showing transparent colony variants
(From 7

rapidly As mentioned earlier, this has been the experience of many workers in relation to chloramphenicol A very striking example of this pehnomenon with erythromycin has been reported by Lepper and his colleagues (1954) from a fever hospital in Chicago The antibiotic was introduced for the treatment of all cases of pertussis or coccal infections other than meningitis in September 1952 Within five months 75 per cent of the staphylococci isolated from hospital personnel were resistant to 100 μgm /ml erythromycin, and 95 per cent of the strains causing respiratory infection in tracheotomized patients were similarly resistant In February 1953 the use of erythromycin was stopped and by June the incidence of erythromycin resistant staphylococci amongst those isolated from hospital personnel had fallen to 29 per cent

Lowbury (1955) found that when the use of chlortetracycline was discontinued in a burns unit the proportion of tetracycline resistant staphylococci in the ward tended to fall at about the same rate as it had increased while the antibiotic was in use Gibson and Thompson (1956) show how the incidence of chlortetracycline resistant strains of *Staph pyogenes* isolated from burns increased over a period of one year during which all burns were treated by local application of this antibiotic, and then decreased when this treatment was discontinued

New Zealand has recently set an example Following the investigations of Beavan and Burry (1956) who state that resistant strains of staphylococci have become a real problem in New Zealand only in the past year since the broad spectrum antibiotics have been available on a doctor s prescription the Minister of Health of New Zealand announced that from 1 February 1956 erythromycin would not be issued to doctors by the Social Security Fund except through hospitals and then only for the treatment of diseases which did not respond to other remedies Moreover hospitals were asked to use the antibiotic carefully

Failing government control which no doubt many doctors would regard as a serious curtailment of their liberty individual hospitals could do much The above mentioned instances of the disappearance of resistant strains following discontinuance of

an antibiotic in a hospital encourage the belief that a careful hospital policy for these drugs could play a major part in the control of antibiotic resistant staphylococcal infection As enough antibiotics are discovered, perhaps, the ones in main use could be changed every six months or so, and even with our present supply it is possible for hospitals to keep at least one anti staphylococcal agent in reserve

Individual doctors in hospitals should be asked to use antibiotics with discrimination, only for the treatment of infections likely to respond, and wherever possible with laboratory control Prophylactic treatment, especially with the broad spectrum antibiotics which eliminate the normal bacterial flora, should be reduced to a minimum Wherever antibiotics are used they should be given in full doses and for adequate periods The emergence of antibiotic resistant staphylococci might also be reduced by the use of two antibacterial agents, rather than one in the treatment of staphylococcal infections, as is now almost universally done in tuberculosis Finally, since prolonged local treatment is the ideal method for encouraging the development of antibiotic resistant variants only antibiotics which are too toxic for systemic administration should be used for local application, and for this bacitracin and neomycin alone, or in combination can be recommended

Control of Cross infection One way to avoid getting antibiotic resistant staphylococcal infection at present is, of course, to keep out of hospitals and both Ludlam (1953) and Edmunds *et al* (1955) have demonstrated the staphylococcal advantages of being born at home A better answer would be to improve hospitals, but this is a major problem and would entail considerable structural alterations Once again however, New Zealand leads the way Thus Pickerill and Pickerill (1954) working in a small hospital for infants undergoing plastic surgery, claim that cross infection has been eliminated by a system of mother nursing, each mother and infant being given a small private ward No doubt few hospitals have the facilities for this, but it is high time we faced the fact that staphylococcal infection is a communicable disease and should not be nursed in the middle of a large general ward

The high incidence of nasal carriers among the nurses in hospitals cannot be ignored As shown by Hare and Thomas (1956), staphylococci from nasal carriers are not spread directly by droplets but follow an indirect route, involving contamination of skin and clothing similar to that of other respiratory organisms (Hare and Mackenzie, 1946) It is, therefore, more important for carriers to wear sterile gowns and gloves than masks when attending to babies or surgical dressings Some investigators have recently claimed a reduction in the incidence of antibiotic resistant staphylococcal infection by the treatment of staff nasal carriers (Gould and Allan, 1954, Rountree *et al* 1956) This may be of value, but, as for all local treatment, antibiotics suitable for systemic use should be avoided Brodie, Kerr and Sommerville (1956) have shown that faecal carriers among patients are another important source of antibiotic resistant staphylococci

Notwithstanding all these carriers the most important measure for avoiding hospital infection with *Staph pyogenes* is to keep all septic cases out of clean wards and theatres and the present policy of nursing patients after clean operations side by side with open staphylococcal infections, and operating on clean and septic cases in the same theatres cannot be too strongly condemned Until a hospital policy of segregating septic cases has been introduced, other measures are unlikely to have much effect

REFERENCES

ABRAHAM E P CHAIN E FLETCHER C M FLOREY A. W GARDNER A D HEATLEY N G and JENNINGS M A (1941) Further observations on penicillin *Lancet* **ii** 177

ALDER V G GILLESPIE W A. and THOMPSON M E M (1955) Virulence and phage patterns of antibiotic resistant staphylococci in hospital *J Path Bact* **70** 503

ANDERSON E S and WILLIAMS R E O (1956) Bacteriophage typing of enteric pathogens and staphylococci and its use in epidemiology *J clin Path* **9** 94

ANTTONEN V M (1955) The resistance of staphylococci to antibiotics and sulfa drugs *Ann Med exper et Biol Fenn* **33** 145

AUSTRIAN R. (1952) Bacterial transformation reactions *Bact Rev* **16** 31

BARBER M (1947a) Coagulase positive staphylococci resistant to penicillin *J Path Bact* 59 373

BARBER, M (1947b) Staphylococcal infection due to penicillin resistant strains *Brit med J* ii, 863

BARBER M (1949) The incidence of penicillin sensitive variant colonies in penicillinase producing strains of *Staph pyogenes J gen Microbiol* 3 274

BARBER M (1953a) The effect of serial passage in other antibiotics on penicillinase producing staphylococci *J gen Microbiol* 8 104

BARBER M (1953b) Penicillin resistant and dependent staphylococcal variants *J gen Microbiol* 8 111

BARBER, M (1953c) Antibiotic resistant staphylococcal variants Adaptation in Micro organisms (The third Symposium of the Society for General Microbiology) Cambridge University Press pp 235–49

BARBER M (1957) In the press

BARBER M and BURSTON J (1955) Antibiotic resistant staphylococcal infection A study of antibiotic sensitivity in relation to bacteriophage types *Lancet* ii 578

BARBER M HAYHOE F G J and WHITEHEAD J E M (1949) Penicillin resistant staphylococcal infection in a Maternity hospital *Lancet* ii 1120

BARBER M and ROZWADOWSKA DOWZENKO M (1948) Infection by penicillin resistant staphylococci *Lancet* ii 641

BARBER M and WHITEHEAD J E M (1949) Bacteriophage types in penicillin resistant staphylococcal infection *Brit med J* ii 565

BARBER M WILSON B D R RIPPON J E and WILLIAMS R E O (1953) Spread of *Staphylococcus aureus* in a maternity department in the absence of severe sepsis *J Obstet Gynaec* 60 476

BARROW G I (1955) Clinical and bacteriological aspects of Impetigo contagiosa *J Hyg Camb* 53 495

BEAVAN D W and BURRY A F (1956) Staphylococcal pneumonia in the newborn An epidemic with eight fatal cases *Lancet* ii 211

BLAIR J E CARR M and BUCHANAN J (1946) The action of penicillin on staphylococci *J Immunol* 52 281

BLOWER R MASON G A WALLACE K R and WALTON M (1955) Control of wound infection in a Thoracic Surgery unit *Lancet* ii 786

BONDI A KORNBLUM J and DE SAINT PHALLE M (1953) Isolation of penicillin susceptible mutants from penicillinase producing strains of *Micrococcus pyogenes Proc Soc exp Biol N Y* 83 527

BRODIE J JAMIESON W and SOMMERVILLE J (1955) Complications of oxytetracycline and tetracycline therapy related to a defined type of resistant staphylococcus *Lancet* ii 223

BRODIE J KERR M R and SOMMERVILLE T (1956) The hospital staphylococcus A comparison of nasal and faecal carrier states *Lancet* i 19

BRODIE J SOMMERVILLE T and WILSON S G F (1956) Coagulase positive staphylococci A serial survey for nasal carriers during the first six months of nursing training *Brit med J* i 667

CAIRNS H J F and SUMMERS G A C (1950) Penicillin resistant staphylococci Incidence in relation to length of stay in hospital *Lancet* **i** 446

CHABBERT Y TERRIAL G and SCHUTZENBERGER M P (1953) Development of the sensitivity to antibiotics of organisms isolated in Paris 1949–1952 *Ann de l'Inst Past* **84** 952

CHANDLER C A DAVIDSON V Z LONG P H and MONNIER J J (1951) Studies on resistance of staphylococci to penicillin The production of penicillinase and its inhibition by the action of aureomycin *Johns Hopk Hosp Bull* **89** 81

CLARKE, S K R. DALGLEISH P G and GILLESPIE W A (1952) Hospital Cross Infections with staphylococci resistant to several antibiotics *Lancet* **i** 1132

CLARKE S K R DALGLEISH P G PARRY E W and GILLESPIE W A (1954) Cross Infection with penicillin resistant *Staph aureus* Effect of oiling floor and bed-clothes in a surgical ward *Lancet* **ii** 211

COLBECK J C (1949) An extensive outbreak of staphylococcal infection in maternity units *Canad med Ass J* **61** 557

DEAN A C R and HINSHELWOOD C (1953) Observations on bacterial adaptation (The third Symposium of the Society for General Microbiology) Cambridge University Press pp 21–39

DEARING W H and HEILMAN F R (1953) Micrococcic (staphylococcal) enteritis as a complication of antibiotic therapy its response to erythromycin *Proc Staff Meet Mayo Clin* **28** 121

DEMEREC M (1945) Production of staphylococcus strains resistant to various concentrations of penicillin *Proc Nat Acad Sci* **31** 215

DEMEREC M (1948) Origin of bacterial resistance to antibiotics *J Bact* **56** 63

DISNEY M E WOLFF J and WOOD B S B (1956) Staphylococcal pneumonia in infacts *Lancet* **i** 767

DOLMAN C E (1933) Treatment of local staphylococcic infections with staphylococcus toxoid *J Amer med Ass* **100** 1007

DOLMAN C E (1935) Clinical uses of staphylococcus toxoid *Lancet* **i** 306

DUBOS R J (1955) The micro-environment of inflammation or Metchnikoff revisited *Lancet* **ii** 1

DUBOS R J SMITH J MACLEAN and SCHAEDLER R W (1955) Metabolic disturbances and infection *Proc Roy Soc Med* **48** 911

EDMUNDS P N ELIAS JONES T F FORFAR J O and BALF C L (1955) Pathogenic staphylococci in the environment of the newborn infant *Brit med J* **i** 990

EHRLICH P (1908) The experimental researches on specific therapeutics London The Harben Lectures 1907

EL GHOROURY A A A (1954) Incidence of staphylococci resistant to penicillin and streptomycin *J Egypt med Ass* **37** 49

ERIKSEN K R. (1952) Hospital Infection with staphylococci resistant to penicillin and other antibiotics *Ugeskr laeg* **114** 1607

BARBER M (1947a) Coagulase positive staphylococci resistant to penicillin *J Path Bact* 59 373

BARBER M (1947b) Staphylococcal infection due to penicillin resistant strains *Brit med J* ii 863

BARBER M (1949) The incidence of penicillin sensitive variant colonies in penicillinase producing strains of *Staph pyogenes* *J gen Microbiol* 3 274

BARBER, M (1953a) The effect of serial passage in other antibiotics on penicillinase producing staphylococci *J gen Microbiol* 8 104

BARBER M (1953b) Penicillin resistant and dependent staphylococcal variants *J gen Microbiol* 8 111

BARBER M (1953c) Antibiotic resistant staphylococcal variants Adaptation in Micro organisms (The third Symposium of the Society for General Microbiology) Cambridge University Press pp 235–49

BARBER M (1957) In the press

BARBER, M and BURSTON J (1955) Antibiotic resistant staphylococcal infection A study of antibiotic sensitivity in relation to bacteriophage types *Lancet* ii 578

BARBER, M, HAYHOE F C J and WHITEHEAD J E M (1949) Penicillin resistant staphylococcal infection in a Maternity hospital *Lancet* ii 1120

BARBER M and ROZWADOWSKA DOWZENKO M (1948) Infection by penicillin resistant staphylococci *Lancet* ii, 641

BARBER M and WHITEHEAD J E M (1949) Bacteriophage types in penicillin resistant staphylococcal infection *Brit med J* ii 565

BARBER M WILSON B D R RIPPON J E and WILLIAMS R E O (1953) Spread of *Staphylococcus aureus* in a maternity department in the absence of severe sepsis *J Obstet Gynaec* 60 476

BARROW, G I (1955) Clinical and bacteriological aspects of Impetigo contagiosa *J Hyg, Camb* 53 495

BEAVAN D W and BURRY A F (1956) Staphylococcal pneumonia in the newborn An epidemic with eight fatal cases *Lancet* ii 211

BLAIR J E CARR M and BUCHANAN J (1946) The action of penicillin on staphylococci *J Immunol* 52 281

BLOWER R MASON G A WALLACE K R and WALTON M (1955) Control of wound infection in a Thoracic Surgery unit *Lancet*, ii 786

BONDI A KORNBLUM J and DE SAINT PHALLE M (1953) Isolation of penicillin susceptible mutants from penicillinase producing strains of *Micrococcus pyogenes* *Proc Soc exp Biol N Y* 83 527

BRODIE J JAMIESON W and SOMMERVILLE J (1955) Complications of oxytetracycline and tetracycline therapy related to a defined type of resistant staphylococcus *Lancet* ii 223

BRODIE J KERR M R and SOMMERVILLE T (1956) The hospital staphylococcus A comparison of nasal and faecal carrier states *Lancet* i 19

BRODIE J SOMMERVILLE T and WILSON, S G F (1956) Coagulase positive staphylococci A serial survey for nasal carriers during the first six months of nursing training *Brit med J* i 667

CAIRNS H J F and SUMMERS G A C (1950) Penicillin resistant staphylococci Incidence in relation to length of stay in hospital *Lancet* i 446

CHABBERT Y TERRIAL G and SCHUTZENBERGER M P (1953) Development of the sensitivity to antibiotics of organisms isolated in Paris 1949–1952 *Ann de l Inst Past* 84 952

CHANDLER C A DAVIDSON V Z LONG P H and MONNIER J J (1951) Studies on resistance of staphylococci to penicillin The production of penicillinase and its inhibition by the action of aureomycin *Johns Hopk Hosp Bull* 89 81

CLARKE S K R DALGLEISH P G and GILLESPIE W A (1952) Hospital Cross Infections with staphylococci resistant to several antibiotics *Lancet* i 1132

CLARKE S K R DALGLEISH P G PARRY E W and GILLESPIE W A (1954) Cross Infection with penicillin resistant *Staph aureus* Effect of oiling floor and bed-clothes in a surgical ward *Lancet* ii 211

COLBECK J C (1949) An extensive outbreak of staphylococcal infection in maternity units *Canad med Ass J* 61 557

DEAN A C R and HINSHELWOOD C (1953) Observations on bacterial adaptation (The third Symposium of the Society for General Microbiology) Cambridge University Press pp 21–39

DEARING W H and HEILMAN F R (1953) Micrococcic (staphylococcal) enteritis as a complication of antibiotic therapy its response to erythromycin *Proc Staff Meet Mayo Clin* 28 121

DEMEREC M (1945) Production of staphylococcus strains resistant to various concentrations of penicillin *Proc Nat Acad Sci* 31 215

DEMEREC M (1948) Origin of bacterial resistance to antibiotics *J Bact* 56 63

DISNEY M E WOLFF J and WOOD B S B (1956) Staphylococcal pneumonia in infacts *Lancet* i 767

DOLMAN C E (1933) Treatment of local staphylococcic infections with staphylococcus toxoid *J Amer med Ass* 100 1007

DOLMAN C E (1935) Clinical uses of staphylococcus toxoid *Lancet* i 306

DUBOS R J (1955) The micro environment of inflammation or Metchnikoff revisited *Lancet* ii 1

DUBOS R J SMITH J MACLEAN and SCHAEDLER R W (1955) Metabolic disturbances and infection *Proc Roy Soc Med* 48 911

EDMUNDS P N ELIAS-JONES T F FORFAR J O and BALF C L (1955) Pathogenic staphylococci in the environment of the newborn infant *Brit med J* i 990

EHRLICH P (1908) The experimental researches on specific therapeutics London The Harben Lectures 1907

EL GHOROURY A A A (1954) Incidence of staphylococci resistant to penicillin and streptomycin *J Egypt med Ass* 37 49

ERIKSEN K R (1952) Hospital Infection with staphylococci resistant to penicillin and other antibiotics *Ugeskr laeg* 114 1607

EVANS A D and EVANS M (1956) Staphylococcal infection of the lower respiratory tract in adults with influenza *Lancet* i 771

FAIRBROTHER R W (1956) Mixed staphylococcal infections The development of penicillin resistant strains *Lancet* i 716

FAIRBROTHER R W PARKER L and EATON B R (1954) The stability of penicillinase producing strains of *Staphylococcus aureus* *J gen Microbiol* 10 309

FINLAND M (1951) The present status of antibiotics in bacterial infections *Bull N Y Acad Med* 27 199

FINLAND M (1955) Changing patterns of resistance of certain common pathogenic bacteria to antimicrobial agents *New Engl J Med* 252 570

FINLAND M COLLINS H S and PAINE T F Jr (1948) Aureomycin a new antibiotic Results of laboratory studies and clinical use in 100 cases of bacterial infection *J Amer med Ass* 138 946

FINNEY J M T (1893) Gastroenterostomy for cicatrizing ulcer of the pylorus *Bull Johns Hopk Hosp* 4 53

FORBES G B (1949) Infection with penicillin resistant staphylococci in hospital and general practice *Brit med J* ii 569

FORFAR J O BALF C L ELIAS JONES T F and EDMUNDS P N (1953) Staphylococcal infection of the newborn *Brit med J* ii 170

FOUACES J and LUTZ A (1953) Type bacteriophage des staphylocoques pathogenes secreteurs de penicillinase *Ann Inst Pasteur* 85 378

FOWLER B J (1955) Post operative staphylococcal enterocolitis during antibiotic therapy *Brit med J* i 1313

FRAME H F and SHORT D W (1955) Drug induced enteritis *Lancet* i 434

FUSILLO M H ROERIG R N METZGER J F and ERNST K F (1954) Phage typing of antibiotic resistant staphylococci *Amer J pub Health* 44 317

GALE E F and RODWELL A W (1949) The assimilation of amino-acids by bacteria 7 The nature of resistance to penicillin in *Staphylococcus aureus* *J gen Microbiol* 3 127

GALE E F and TAYLOR E S (1947) The assimilation of amino acids by bacteria 2 The action of tyrocidin and some detergent substances in releasing amino-acids from the internal environment of *Strep faecalis* *J gen Microbiol* 1 77

GARDNER D L (1953) Aureomycin resistant staphylococcal enterocolitis *Lancet* ii 1236

GARROD L P and WATERWORTH P M (1956) Behaviour of some new antistaphylococcal antibiotics *Brit med J* ii 61

GIBSON C D Jr and THOMPSON W C (1956) The response of burn wound staphylococci to alternating programmes of antibiotic therapy *Antib Annual N Y 1955–56*

GILSON B St C and PARKER R F (1948) Staphylococcal penicillinase characteristic of the enzyme and its distribution *J Bact* 55 801

GLASMAN M G (1953) Studies of resistant strains of staphylococci *Microbiologica* (U S S.R) no 2 p 53

GOULD A. C. and ALLAN W S A (1954) *Staphylococcus pyogenes* cross infection Prevention by treatment of carriers *Lancet* ii 988

HARE R and MACKENZIE D M (1946) The source and transmission of nasopharyngeal infections due to certain bacteria and viruses *Brit med J* i 865

HARE R. and THOMAS C. G A (1956) The transmission of *Staphylococcus aureus* *Brit med J* ii, 840

HAY P and McKENZIE P (1954) Side effects of oxytetracycline therapy *Lancet* i 945

HAZEN L N JACKSON C G CHANG SHIH MAN PLACE E H and FINLAND M (1951) Antibiotic treatment of pertussis *J Pediat* 39 1

HOBSON D (1954) Activity of erythromycin against *Staphylococcus aureus* *Brit med J* i 236

HOWIE J W DUNCAN I B R and MACKIE L. M (1953) Growth of *Clostridium welchii* in the stomach after partial gastrectomy *Lancet* ii 1018

JACKSON G G HAIGHT T H KASS E H WOMACK C. R. GOCKE M and FINLAND M (1951) Terramycin therapy of Pneumonia clinical and bacteriological studies in 91 cases *Ann int Med* 35 1175

JACKSON G G LEPPER M H and DOWLING H F (1954) Bacteriophage typing of staphylococci III Relationship to antibiotic sensitivity and resistance *J Lab clin Med* (Chicago) 44 41

JANBON L BERTRAND J ROUX J and SALVAING J (1952) Super infections staphylococciques de l intestine et accidents choleriformes de la terramycine *Bull Acad nat Med Paris* 136 59

KIRBY W M M and AHERN J J (1953) Changing pattern of resistance of staphylococci to antibiotics *Antib & Chemoth* 3 831

KLECKNER M S BARGEN J A. and BAGGENSTOSS A H (1950) Pseudomembranous enterocolitis clinicopathologic study of fourteen cases in which the disease was not preceded by an operation *Gastroenterology* 21 212

KOSSIAKOFF M G (1887) The property possessed by microbes of adapting themselves to antiseptic media *Ann Inst Pasteur* 1 465

LEDERBERG J and LEDERBERG E M (1952) Replica plating and indirect selection of bacterial mutants *J Bact* 63 399

LEPPER M H MOULTON B DOWLING H F JACKSON G G and KOFMAN S (1954) Epidemiology of erythromycin resistant staphylococci in a hospital population Effect on therapeutic activity of erythromycin *Antib Annual 1953-54* p 308

LOMINSKI I and ROBERTS G B S (1946) A substance in human serum inhibiting coagulase *J Path Bact* 58 187

LOVE B D Jr WRIGHT S S PURCELL E M MOU T W and FINLAND M (1954) Antibacterial action of tetracycline comparison with oxytetracycline and chlortetracycline *Proc Soc exp Biol N Y* 85 25

LOWBURY E J L (1955) Cross Infection of wounds with antibiotic resistant organisms *Lancet* 1 985
LOWBURY E J L TOPLEY E and HOOD A M (1952) Chemotherapy for *Staphylococcus aureus* in burns *Lancet* 1 1036
LUDLAM G B (1953) Incidence and penicillin sensitivity of *Staphylococcus aureus* in the nose in infants and their mothers *J Hyg Camb* 51 64
LURIA S E (1946) A test for penicillin sensitivity and resistance in the staphylococcus *Proc Soc exp Biol N Y* 61 46
LURIA S E and DELBRUCK M (1943) Mutations of bacteria from virus sensitivity to virus resistance *Genetics* 28 491
LYONS C (1937) Antibacterial immunity to *Staph pyogenes Brit J exp Path* 18 411
MONNIER J J and SCHOENBACH E B (1951) The resultant sensitivity of micro organisms to various antibiotics after induced resistance to each of these agents *Antib & Chemoth* 1 472
NICHOLS D R and NEEDHAM G M (1949) Aureomycin in the treatment of penicillin resistant staphylococcic bacteremia *Proc Mayo Clin* 24 309
NICHOLS R L and FINLAND M (1956) Novobiocin a limited bacteriological and clinical study of its use in forty five patients *Antib Med* 2 241
OGINSKY E L SMITH P H and UMBREIT W W (1949) The action of streptomycin I The nature of the reaction inhibited *J Bact* 58 747
PAINE T F Jr COLLINS H S and FINLAND M (1948) Bacteriologic studies on aureomycin *J Bact* 56 489
PARISH H J O MEARA R A Q and CLARK W H M (1934) The clinical investigation of staphylococcal toxin toxoid and antitoxin *Lancet* 1 1054
PARKER M T TOMLINSON A J H and WILLIAMS R E O (1955) Impetigo contagiosa The association of certain types of *Staphylococcus aureus* and of *Streptococcus pyogenes* with superficial skin infection *J Hyg Camb* 53 458
PENNER A and BERNHAIM A I (1939) Acute postoperative enterocolitis A study on the pathologic nature of shock *Arch Path* 27 966
PETTET J D BAGGENSTOSS A H DEARING W H and JUDD E S Jr (1954) *Postoperative pseudomembranous enterocolitis Surg Gynec Obstet* 98 546
PICKERILL C M and PICKERILL H P (1954) Elimination of hospital cross infection in children Nursing by mother *Lancet* 1 425
PRISSICK F H (1953) Antibiotic resistant staphylococci and related infections *Amer J med Sci* 225 299
RAKE G MCKEE C M HAMRE D M and HOUCK C L (1944) Studies on Penicillin II Observations on therapeutic activity and toxicity *J Immunol* 48 271
REES E G SHOOTER R A and SHAWE G D H (1955) Sensitivity to five antibiotics of a further 200 strains of *Staph pyogenes* isolated from out patients *Brit med J* 1 1409

REIDEL J (1902) Ueber darmdiphtherie nach schweren operationem bei sehr geschwachten kranken *Dtsch Z Chir* 67 402

ROODYN L (1954) Staphylococcal infections in general practice *Brit med J* ii 1322

ROODYN L (1956) Staphylococcal infections in general practice *Proc Roy Soc Med* 49 263

ROUNTREE P M (1956) Staphylococci harboured by people in Western Highlands of New Guinea *Lancet* i 719

ROUNTREE P M and BARBOUR R G H (1951) Nasal carrier rates of *Staphylococcus pyogenes* in hospital nurses *J Path Bact* 63 313

ROUNTREE P M and FREEMAN B M (1955) Infections caused by a particular phage type of *Staphylococcus aureus* *Med J Aus* ii 157

ROUNTREE P M HESELTINE M RHEUBEN J and SHERMAN R P (1956) Control of staphylococcal infection of the newborn by the treatment of nasal carriers in the staff *Med J Aus* i 528

ROUNTREE P M and RHEUBEN J (1956) Penicillin resistant staphylococci in the general population *Med J Aus* i 39

ROUNTREE P M and THOMSON E F (1949) Incidence of penicillin resistant and streptomycin resistant staphylococci in a hospital *Lancet* ii ,01

ROUNTREE P M and THOMSON E F (1952) Incidence of antibiotic resistant staphylococci in a hospital *Lancet* ii 262

SCHATZ A BUGIE E and WAKSMAN S A (1944) Streptomycin a substance exhibiting antibiotic activity against gram positive and gram negative bacteria *Proc Soc exp Biol N Y* 55 66

SMITH G N and WORREL C S (1949) The enzymatic hydrolysis of chloramphenicol (chloromycetin) *Science* 110 297

SMITH G N and WORREL C S (1950) The decomposition of chloramphenicol by micro-organisms *Arch Biochem* 28 232

SMITH P H OGINSKY E L and UMBREIT W W (1949) The action of streptomycin II The metabolic properties of resistant and dependent strains *J Bact* 58 761

SPINK W W (1951) Clinical and biological significance of penicillin resistant staphylococci including observations with streptomycin aureomycin chloramphenicol and terramycin *J Lab clin Med* 27 278

SPINK W W FERRIS V and VIVINO J J (1944) Antibacterial effect of whole blood upon strains of staphylococci sensitive and resistant to penicillin *Proc Soc experim Biol N Y* 55 210

SURGALLA M J and DICK G M (1955) Enterotoxin produced by micrococci from cases of enteritis after antibiotic therapy *J Amer med Ass* 158 649

TERPLAN K (1953) Fatal fulminating staphylococcic gastro enterocolitis with shock like state following antibiotic treatment *Amer J Path* 29 595

THOMSON E F ROUNTREE P M and FREEMAN B F (1956) Observations on the sensitivity to erythromycin of *Staphylococcus aureus* *Antib Annual 1955–56* p 63

VALENTINE F C O and BUTLER E C B (1939) Specific immunity in acute staphylococcal osteomyelitis *Lancet* **i** 973

VALENZUELA P E and VACCARO H (1954) Sensibilidad de 688 cepas de *Micrococcus pyogenes* (staphylococcus) a cinco antibioticos Estudio comparativo de las cepas resistentes aisladas en los anos 1951 y 1953 *Rev Med Chile* **82** 356

VOGELSANG T M (1953) Staphylococcal studies in hospital staffs II Penicillin resistance determinations *Acta pathol microbiol scand* **33** 301

WALLMAN I S GODFREY R C and WATSON J R H (1955) Staphylococcal pneumonia in infancy *Brit med J* **ii** 1423

WALLMARK G (1954a) Bacteriophage typing of *Staphylococcus aureus pyogenes* II An analysis of strains isolated from purulent lesions in hospitalized patients including sensitivity tests to antibiotics *Acta pathol microbiol scand* **34** 497

WALLMARK G (1954b) Bacteriophage typing of *Staphylococcus aureus pyogenes* III An analysis of strains isolated from purulent lesions in out patients including sensitivity tests to antibiotics *Acta pathol microbiol scand* **34** 577

WHITBY L E H (1936) The treatment of staphylococcal skin lesions with toxoid *Lancet* **i** 1454

WILLIAMS E (1954) Staphylococcal pseudomembranous enterocolitis complicating treatment with aureomycin *Lancet* **ii** 999

WILLIAMS M R and PULLAN J M (1953) Necrotising enteritis following gastric surgery *Lancet* **ii** 1013

WILLIAMS R E O RIPPON J E and DOWSETT L M (1953) Bacteriophage typing of strains of *Staphylococcus aureus* from various sources *Lancet* **i** 510

WISE R CRANNY C and SPINK W W (1954) Epidemiological studies on antibiotic resistant staphylococci *Bull Univ Minnesota Hosps* **26** 174

WISE R I VOIGT A E COLLIN M V CRANNY C L (1955) Origin of Erythromycin Resistant strains of *Micrococcus Pyogenes* in Infections Bacteriophage types and *in vitro* resistance of cultures to antibiotics *Arch intern Med* **95** 419

WOMACK C R JACKSON C G GOCKE T M KASS E H HAIGHT T H and FINLAND M (1952) Terramycin therapy of urinary tract infections *Arch int Med* **89** 240

XX

Poliomyelitis

H J SEDDON

THIS lecture is concerned with some recent work on certain aspects of paralytic poliomyelitis Much of it has been carried out at the Royal National Orthopaedic Hospital and the Institute of Orthopaedics but here and there I shall refer to relevant work done elsewhere

The extremely variable vulnerability of the central nervous system in poliomyelitis is a most mysterious affair though we are beginning to find a few clues such as the harmful influence of injury (Levinson Milzer and Lewin 1945, Grossiord, Wimphen and Desproges Gotteron, 1953) including certain types of injections (Medical Research Council 1956) the role of exercise (Russell 1947 Horstmann, 1950) though recently doubt has been thrown on its importance (Sutherland 1956) and the evil effects of tonsillectomy and pregnancy All these factors tend to increase the susceptibility of the nervous system during the incubation period of the disease

There can be an increased local vulnerabilty too Paralysis provoked by injection is often predominant in the limb into which the injection was made paralysis following tonsillectomy is frequently bulbar Yet apart from such known localizing factors there is a fairly constant predilection for the lumbo-sacral cord This has been demonstrated repeatedly as in the Malta epidemic of 1943 (Seddon Agius Bernstein and Tunbridge 1945), and various ingenious hypotheses have been put forward to explain it but not one of them carries conviction

Some of our recent work has been focused on an even

narrower aspect of the problem, the detailed anatomical distribution of the lesions in the spinal cord

Cell Stations

Two features of poliomyelitis have long been recognized, though in rather a vague way Although at first sight the distribution of the paralysis appears to follow no particular pattern, those who have studied the disease closely have, in fact, noticed that certain combinations of muscles are affected with some regularity The anterior and posterior tibial muscles are frequently involved together likewise the quadriceps and the flexors of the hip the biceps femoris and the calf muscle The association is by no means one of function It has also been observed that certain muscles or muscle groups are particularly vulnerable, there is a more than average chance that paralysis of them will be permanent Well known examples of this are the tibial muscles and the thenar muscles

Some years ago I discussed these questions with W J W Sharrard and we felt that there must be an anatomical explanation for both phenomena He then embarked on a formidable investigation which is not yet complete (Sharrard, 1955b) We had a large collection of clinical records of cases of poliomyelitis in which muscle power was fairly accurately documented Some of the patients were in a precarious condition on account of more or less severe respiratory paralysis and a few of them died Seven spinal cords were obtained and Sharrard cut serial sections of them and set about counting all the anterior horn cells The work was extremely tedious but the underlying idea was simple enough If only one muscle say the tibialis anterior had been completely knocked out then it was reasonable to suppose that a zone of deficiency in the lumbar anterior horn was due to destruction of the cells innervating the tibialis anterior The catch here however is that small zones of destruction may be found in cords from patients with no clinically demonstrable paralysis (see below) and correlation between paralysis of one muscle and what one finds in the cord involves too much guesswork The reverse state of affairs which in the severely paralysed patient is much more common is far more

informative, the presence of activity in a few muscles in an otherwise totally paralysed limb would be matched by the pre sence of well defined zones of anterior horn cells in a cord from which all other motor cells had been swept away At this stage in the work it was necessary to make one important and rather great assumption, that in a case of poliomyelitis in which there was extensive paralysis there would not be found a more or less extensive pool of surviving motor neurons which for some reason or other were inactive That such dormant neurons might exist had long been a belief underlying a good deal of the practice of physiotherapy, the aim of treatment was to coax these in active neurons back into a state of usefulness It soon became evident that the assumption was justified and that in the patient who had been well treated from the first there was an extremely strict correspondence between the degree of muscle involvement as observed clinically and the number of surviving anterior horn cells

Thus Sharrard was able to use the focal lesions produced by poliomyelitis as a means of determining the site and extent of the cell stations of all the important muscles in the lower limb The root supplies of these muscles had long been known but not the sites of the cell stations themselves He found that the pattern of cell stations and their relationship one with another was constant and this explained the association that I have already mentioned, between paralysis of two or more muscles or muscle groups If one cell station was knocked out by the virus then there was a strong probability that its neighbours would be affected too Sharrard went further He argued that the clinical behaviour of one muscle group might give a clue to the expected behaviour of those associated with it If re covery failed to occur in one group then there was a strong prob ability that it would not occur in muscles innervated by the neighbouring cell stations On the other hand, a flicker of re covery in say the quadriceps might be taken as an indication that there was some hope for its paralysed next door neighbour in the cord the flexors of the hip

The particular vulnerability of certain muscles was similarly explicable The vulnerable muscles were found to have short

cell stations, the less susceptible muscles long ones (Table 1) If the damage done by the poliomyelitis virus could be likened to that caused by a bomb a single house was more likely to be totally destroyed than a long building such as a railway shed

TABLE 1 The relationship between length of cell column and the ratio of paresis to paralysis in muscles in the lower limb (Sharrard 1955b)

Muscle	Ratio of paresis to paralysis	Approximate length of motor cell column (millimetres)
Tibialis anterior	0 47	8
Flexor digitorum longus	0 52	8
Tibialis posterior	0 57	8
Flexor hallucis longus	0 58	8
Extensor hallucis longus	0 89	9
Extensor digitorum longus	1 06	10
Peronei	1 10	10
Calf muscles (triceps surae)	1 26	14
Biceps femoris	2 03	14
Hip abductors	2 78	16
Intrinsic foot muscles	2 87	15
Inner hamstring muscles	2 90	20
Tensor fasciae latae	2 92	16
Quadriceps	3 00	22
Gluteus maximus	3 30	17
Hip adductors	3 50	22
Hip flexors	3 67	23

In the course of this work Sharrard was able to correlate the degrees of paralysis expressed by the system of grading adopted by the Medical Research Council during the war with the proportion of anterior horn cells surviving (Table 2) A remarkable finding was that up to about half of the anterior horn cells in a cell station may be destroyed without any noticeable weakening of the muscle as detected by ordinary clinical examination This explains why it is risky to attempt to localize cell stations in cases where there is minimal paralysis Nothing less than elaborate comparative examination with a spring balance or dynamometer will reveal these deficiencies On the other hand it was surprising that muscles with as few as 10 per cent of their anterior horn cells surviving could still function usefully

An important practical lesson to be learned from this investigation is that the actual paralysis observed at the end of the period of recovery is an accurate expression of the extent of permanent damage There is no anatomical basis for the belief that

TABLE 2 The relationship between muscle power and residual motor cells in the spinal cord (Sharrard 1955b)

Muscle power (M R C scale)	Percentage of residual motor cells
0	0–2
1	2–3
2	3–5
3	5–10
4	10–20
4	20–40
5	over 40

a further return of power can be obtained by somehow charming unused motor neurons back to useful activity

This work has been completed only for the muscles of the lower limb but it will not be long now before a similar map of the cervical cord is also available

THE PROCESS OF RECOVERY

From what I have just said it might be supposed that the lesions in poliomyelitis are of a simple all or none character but it must not be forgotten that in selecting his material Sharrard chose spinal cords in which there was good reason to believe that an end point had been reached All sorts of complicated changes might have taken place before finality was reached

We can be quite clear in our minds about the neurons that have been utterly destroyed The anterior horn cells are rapidly removed by phagocytosis the axons and their myelin sheaths degenerate rapidly there is denervation atrophy of the corresponding muscle fibres and indeed these motor units have ceased to exist

We can be equally sure that many anterior horn cells are only temporarily damaged From observations made on experimental animals and to a lesser extent on human material we

cell stations the less susceptible muscles long ones (Table 1) If the damage done by the poliomyelitis virus could be likened to that caused by a bomb, a single house was more likely to be totally destroyed than a long building such as a railway shed

TABLE 1 The relationship between length of cell column and the ratio of paresis to paralysis in muscles in the lower limb (Sharrard 1955b)

Muscle	Ratio of paresis to paralysis	Approximate length of motor cell column (millimetres)
Tibialis anterior	0 47	8
Flexor digitorum longus	0 52	8
Tibialis posterior	0 57	8
Flexor hallucis longus	0 58	8
Extensor hallucis longus	0 89	9
Extensor digitorum longus	1 06	10
Peronei	1 10	10
Calf muscles (triceps surae)	1 26	14
Biceps femoris	2 03	14
Hip abductors	2 78	16
Intrinsic foot muscles	2 87	15
Inner hamstring muscles	2 90	20
Tensor fasciae latae	2 92	16
Quadriceps	3 00	22
Gluteus maximus	3 30	17
Hip adductors	3 50	22
Hip flexors	3 67	23

In the course of this work Sharrard was able to correlate the degrees of paralysis expressed by the system of grading adopted by the Medical Research Council during the war with the proportion of anterior horn cells surviving (Table 2) A remarkable finding was that up to about half of the anterior horn cells in a cell station may be destroyed without any noticeable weakening of the muscle as detected by ordinary clinical examination This explains why it is risky to attempt to localize cell stations in cases where there is minimal paralysis Nothing less than elaborate comparative examination with a spring balance or dynamometer will reveal these deficiencies On the other hand it was surprising that muscles with as few as 10 per cent of their anterior horn cells surviving could still function usefully

excited by the electrodes In a case where the superficial part of the muscle is denervated but the deeper fibres still retain their connection with anterior horn cells the electrical reactions are misleading Although such a muscle would give a reaction of denervation, more or less recovery in the inaccessible part

TABLE 3 Electrical reactions

At three months		
	Reactions normal	—Recovery in 83%
		—No recovery in 17%
	Reaction of denervation	—Recovery in 30%
		—No recovery in 70%
At four months		
	Reactions normal	—Recovery in 100%
		—No recovery in 0 %
	Reaction of denervation	—Recovery in 12%
		—No recovery in 88

would be apparent with the passage of time Brooks (1953) got over this difficulty by stimulating nerve trunks such as the sciatic radial, and ulnar and in this way was able to pick up all the muscle which had retained its anterior horn cell connections Even so there are still a few strange discrepancies (Table 4) Of the muscles that completely failed to respond on stimulation of the corresponding nerve trunk 10 per cent showed some recovery subsequently, which may indicate that as a rare event reversible Wallerian degeneration can occur By contrast of the muscles that responded on stimulation of the nerve trunk 22 per cent failed to recover This may mean that

know that the chief anatomical change is the temporary disappearance of the chromatin (Nissl granules) from the cytoplasm The axis cylinders remain intact and there is no denervation atrophy of the corresponding muscle fibres In this state of affairs *it is possible to make the muscle contract by stimulation of the appropriate nerve trunk, the muscle, of course,* shows no reaction of denervation The anterior horn cell recovers its normal structure and within a comparatively short time begins to work again and voluntary power returns

However, it is conceivable that there is an intermediate state —which Einarson (1949) claims to have demonstrated histologically—in which *the damage to the anterior horn cell is more severe though still ultimately recoverable This state of affairs* might be accompanied by Wallerian degeneration of the axis cylinder Let us examine the implications of this hypothesis In the first place we should find that during the period of Wallerian degeneration muscles would show a reaction of denervation but when recovery occurred the electrical excitability would return to normal *Likewise a nerve trunk would be inexcitable and later regain its excitability* This is what happens after division and repair of a peripheral nerve The question has been investigated by my colleague D M Brooks He has failed to demonstrate any regular reversibility of the reaction of denervation If muscles retain their electrical excitability their cell stations will recover and the muscles will regain *their activity Even more important, if a reaction of denervation* is found it is generally permanent There is no reversal, once the axon has degenerated as a result of damage to the anterior horn cell it will never recover

However Brooks had two reasons for being dissatisfied with the reliability of electrical reactions of muscle as a method *for settling this point After peripheral nerve injury the reaction of degeneration is fully developed by* the twenty first day after poliomyelitis for some reason quite unknown, it takes longer and as Table 3 shows a firm prognosis cannot be made within three months Furthermore in poliomyelitis the electrical examination of muscles has a certain intrinsic limitation Only the superficial part of a *muscle of some bulk* such as the calf, is

to the most peripheral part of a nerve and is of a constant extent whatever the length of the nerve fibre But there is more to it than this If axonal regeneration were a significant factor then the rate of recovery would be governed by the rate of regenera tion which at best, is a slow process It can be seen from Figure

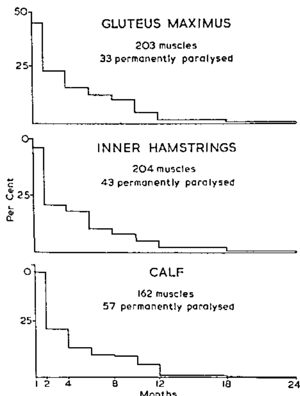


FIG 1 Percentage of muscles showing unit increase in power in each period of two months The general tempo and pattern of recovery are the same irrespective of the d stance of a muscle from the spinal cord (from Sharrard 1955a) (Reproduced from *British Surgical Progress* 1954 by permission of the Editors and Messrs Butterworth and Co)

1 that the greater part of recovery takes place early at a rate far greater than would be compatible with axonal regenera tion over more than a very short distance

the damage in the cord responsible for the paralysis of these muscles was proximal to the anterior horn cells and had knocked out internuncial neurons

There is further evidence against axonal regeneration After severance and repair of a peripheral nerve the muscles regain

TABLE 4

NERVE STIMULATION	Positive response	Recovery in 78%
		No recovery in 22%
	No response	Recovery in 10%
		No recovery in 90%

power in their order of innervation The more proximal ones recover before the more distal and from this it has been possible to work out the rate of regeneration in man quite accurately (Seddon, Medawar and Smith, 1943) It is of the order of about one millimetre a day If axonal regeneration plays any significant part in recovery from poliomyelitis, it is reasonable to expect that in a substantial proportion of cases, the proximal muscles would recover earlier than the distal ones Sharrard (1955a) went into this question He found that although the frequency of recovery of different muscle groups was not the same—for reasons that I have already given—the pattern of recovery was remarkably constant Whatever the distance of a muscle from the spinal cord its rate of recovery was almost always the same (Figure 1) This could not possibly be explained in terms of axonal regeneration unless we made the assumption —though it is not fantastic—that axonal degeneration is limited

It was shown by Harry (1938) and Sharrard (1955a) that the return of muscle power may continue at a very slow rate for up to two years after the acute disease (Figure 1) It is supposed that some of this late recovery is due to the hypertrophy of surviving muscle fibres Certainly, the most remarkable hypertrophy can appear after poliomyelitis and it may be more important than we have realized

Treatment of the Paralysis

Bodian s (1949) demonstration that the cycle of changes in the anterior horn cells of a monkey suffering from poliomyelitis is complete by the thirty fifth day has had an unfortunate effect on policy (in some quarters) in the treatment of poliomyelitis Bodian s observations which are unassailable are limited to the *appearance* of the cells by the end of the fifth week they have either been totally destroyed or they have returned to normal It has, therefore been inferred that at the end of this period, indeed if not earlier, the clinician is dealing with a stable situation one in which the most has to be made of the motor units that have been spared The patient is encouraged to do as much as he can he is stimulated to work hard so that he will make the best possible use of surviving neurons before he has forgotten how to perform the movements that have been lost during the acute phase of the disease Such early activity would presumably encourage muscle hypertrophy too Pushing activity to the point of fatigue would be beneficial rather than harmful

All this is based on the assumption that an anterior horn cell that looks normal in a stained preparation is capable of functioning normally Yet it is within the experience of everyone who ha had much to do with the treatment of poliomyelitis that undue activity in the first few months after the acute attack may cause regression It is therefore, wiser to make haste slowly and it is likely—though we have not yet proved it satisfactorily—that the gentle regime of exercise, as commonly practised in this country is the best

Yet it is possible to go too far in the other direction and here we must consider the function of splints It is not many years

It may well be asked what is the practical importance of this discussion In the treatment of peripheral nerve injuries, we know exactly where we are We can calculate fairly accurately when recovery can be expected in this or that group of muscles and we can control the degree of denervation atrophy by daily galvanic stimulation, which has the effect of maintaining muscle fibres at something approaching their normal bulk (Jackson, 1945) and preventing the harmful and ultimately irreversible interstitial fibrosis which is a secondary feature of denervation atrophy This treatment is continued until a paralysed muscle is beginning to show signs of activity Electrical stimulation of muscle has been used most assiduously in the treatment of poliomyelitis on the assumption that the prevention of denervation atrophy is important It is—if axonal regeneration occurs But in the absence of such a process it is a waste of time and money and inflicts unnecessary suffering on the patient It is not a comfortable form of therapy Thus this evidence against axonal regeneration has the effect of simplifying the treatment of poliomyelitis in eliminating electrotherapy as a useful form of treatment

Changes in Muscle

In poliomyelitis the atrophy that occurs in denervated muscle is no different from that seen after severance of a peripheral nerve except in its distribution If as is often the case, a muscle is only partially denervated patches of shrunken muscle fibres are intermingled with zones in which the fibres are normal It may be that this has an important bearing on recovery Van Harreveld (1945) and Weiss and Edds (1946) have found experimentally that partially denervated muscle may show an appreciable and quite rapid augmentation of power It is suggested that this might be due to the sprouting of axons from intact motor units which make connection with the empty end plates of adjacent denervated muscle fibres and so restore them to functional activity Morris (1953) has demonstrated very beautifully that this process occurs in the rabbit between the fourth and sixtieth day after partial denervation It may occur in poliomyelitis but there is as yet no proof of it

paralysed abductors of the hip and particularly the tensor fascia latae may become shortened if as is often the case the patient lies with the limbs abducted I have already mentioned that progressive interstitial fibrosis is a concomitant of the atrophy of denervated muscle and it is this tissue which shortens if given the opportunity However, there is a factor which may be peculiar to poliomyelitis, contractures can develop extremely rapidly, within the first week or two, and more particularly if the limb is painful This is long before interstitial fibrosis begins It is unlikely that we shall understand the pathology of this disorder until we know more about the causes of pain in early poliomyelitis and their possible connection with early contracture Fortunately the condition is usually preventable if passive movements are carried out persistently from the first

Once a contracture has occurred a whole sequence of untoward events can follow the joint capsule shortens on the side of the contracture and lengthens often permanently on the side where the muscles are overstretched Paralytic dislocation of the joint may follow Both the ligamentous laxity and paralytic dislocation can be greatly aggravated by the effect of gravity—seen at its worst in the lower limb However it should not be supposed that contracture alone is synonymous with disabling deformity Some contractures are compatible with good function and may indeed enhance it for a patient with weakness of the flexors of the elbow a flexion contracture of that joint is a positive advantage Slight shortening of the Achilles tendon in a patient with paralysis of the quadriceps increases the stability of the extended knee What is more deformity can occur in the absence of any contracture However vigorous the efforts that have been made to keep muscles the right length the clinician may be defeated by a gross disturbance of muscle balance If for example the evertors of the foot are completely paralysed and the tibial muscles are more or less normal the foot will inevitably be pulled into inversion No amount of manipulation or splinting will resist this great force Thus loss of muscle balance is not only the prime cause of contracture but is the chief cause of deformity even in the absence of contracture

since enforced rest of weak or paralysed muscles was widely practised A muscle at rest may be in the relaxed, the neutral, or the stretched position For the extensors of the wrist these positions are, respectively extension of the wrist, the neutral position with the carpus in line with the forearm, and the flexed, drop wrist position It is still commonly supposed that relaxation of a weak or paralysed muscle is beneficial, if only because it is quite certain that the opposite, the stretched position, is harmful Yet it may well be that the neutral position is the right one This was investigated by Sharrard and Knowelden (1956) in relation to splinting in abduction for paralysis of the shoulder muscles, in particular the deltoid Here the position of abduction is one of relaxation, the neutral position is with the arm lying by the side, there is no overstretched position They made an exact comparison of recovery in cases that were treated uniformly except that some were splinted in abduction whereas others were not The latter showed a slightly higher rate of recovery Thus much of the expensive and uncomfortable apparatus that we have been using for so long can be discarded

Deformity

As a cause of deformity poliomyelitis has few equals, and formerly a great deal of an orthopaedic surgeon s time was spent in trying to straighten out gross distortions of the human frame Recently the intense interest in respiratory paralysis and the treatment of the early phases of the disease have resulted in less attention being paid to the genesis and prevention of deformity But the danger is always there, and even today severe and often entirely preventable deformity occurs in patients almost on our doorsteps

The trouble begins in the muscles If a muscle is paralysed and its opponents are active the latter will shorten rapidly and a contracture will develop unless vigorous attempts are made to maintain the full length of the active muscle This is well known It is less clearly recognized that a paralysed muscle can develop a contracture if it is splinted in a relaxed and therefore shortened position The best known example in the upper limb is the extensor communis digitorum In the lower limb

body this can be done There has been much interest during the last few years in the prevention of paralytic deformities of the foot by early tendon transplantation and this subject has been studied by Mortens and Pilcher (1956) They have worked out which of the numerous transplantations that have been attempted can be relied on to give satisfactory results The tendons of one or more of the muscles producing a deformity are transplanted either to a site where they will have an opposite and one hopes correct pull or to a neutral position If both peronei are paralysed and the tibial muscles are active the tibialis anterior is shifted to the outer border of the foot The principle is simple enough but the technical details are complicated, and unless the greatest judgement is used an operation of this type will fail it may even result in the production of a deformity the opposite of the one it was designed to prevent In certain situations no transplantation is available and even more drastic methods have been used, namely, the denervation of the muscle that is causing the trouble A pathological increase in the arch of the foot due to unopposed action of the short muscles can be alleviated if the operation is performed early enough, by selective denervation of the muscles of the sole of the foot (Garceau and Brahms, 1956)

REFERENCES

BODIAN D (1949) *International Poliomyelitis Conference* J B Lippincott 6–

BROOKS D M (1953) *The Spinal Cord* Ciba Foundation Symposium J and A Churchill London 280

EINARSON L (1949) *Act orth scand* 19 27

GARCEAU G J and BRAHMS M A (1956) *J Bone & Joint Surg* 38a 553

GROSSIORD A WIMPHEN A and DESPROGES GOTTERON R. (1953) *La Semaine des Hop de Paris* 29 255

HARREVELD A VAN (1945) *Amer J Physiol* 144 477

HARRY N M (1938) *Brit med J* 1 164

HORSTMANN D (1950) *J Amer med Assoc* 142 236

JACKSON E C S (1945) *Brain* 68 300

JAMES J I P (1956) *J Bone & Joint Surg* 38b 660

LEVINSON S O MILZER A and LEWIN P (1945) *Amer J Hyg* 42 204

MEDICAL RESEARCH COUNCIL (1956) *Lancet* ii 1223

It is well known that age is the most important factor predisposing to deformity The younger the patient the greater the probability of deformity developing as a consequence of loss of muscle balance This is exemplified by Figure 2 taken from James s (1956) recent paper on paralytic scoliosis In the young

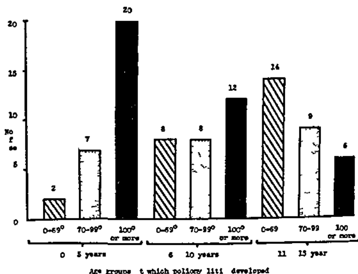


FIG 2 The severity of paralytic curvature of the spine in relation to the age at onset of poliomyelitis (from James 1956) (Reproduced from *J Bone and Joint Surg* by permission of the Editors and Messrs E & S Livingstone Ltd)

there is great plasticity of bones and ligaments Beyond the age when growth ceases the risk of deformity developing is small The consequences of a particular pattern of paralysis are predictable It has long been recognized that certain deformities of the foot are always associated with particular distributions of paralysis Recently James has shown that the same holds good in the trunk Paralysis of the upper intercostals, for example, will give rise to an upper thoracic scoliosis, convex towards the side of the paralysis and characterized by a drooping of the ribs on that side so that they lie almost vertically

The obvious remedy for this state of affairs is to correct the balance that has been disturbed, and in certain parts of the

XXI

The Electron Microscope in the Study of Viruses

KENNETH M SMITH

BEFORE I talk about the electron microscopy of viruses I would like to refresh your memory on some of the chief characteristics of these interesting agents All plant viruses and most of the animal viruses are far below the resolving power of the best optical microscope They cannot be cultivated upon any synthetic cell free medium but can only multiply inside a living susceptible cell All viruses which have been purified and analysed have been shown to be nucleo proteins

The first scientific proof of the existence of such a thing as a virus was made by the Russian botanist Iwanowski in 1892 when he showed that the sap from a tobacco plant infected with mosaic disease was still infectious to healthy tobacco plants after passing a bacterium proof filter candle Four years later two German workers Loeffler and Frosch showed that foot and mouth disease in cattle was also caused by a filter passing agent Now, of course we know that viruses attack every kind of living organism from bacteria to man and over 300 separate viruses are known to attack plants alone

In this short account, I am, I fear less concerned with viruses of medical importance than with those attacking plants and insects since these have been my special study However, once isolated from their respective hosts there is a fairly close super ficial resemblance between the different types of viruses so far as their appearance in the electron microscope is concerned

There is rather a tendency on the part of virologists studying

MORRIS D D B (1953) *J Bone & Joint Surg* **35b** 650
MORTENS J and PILCHER M (1956) *J Bone & Joint Surg* **38b** 633
RUSSELL W R (1947) *Brit med J* **ii** 1023
SEDDON, H J AGIUS T BERNSTEIN H G G and TUNBRIDGE R E (1945) *Quart J Med* **14** 1
SEDDON, H J MEDAWAR P B and SMITH H (1943) *J Physiol* **102** 191
SHARRARD W J W (1955a) *J Bone & Joint Surg* **37b** 63
SHARRARD W J W (1955b) *J Bone & Joint Surg* **37b** 540
SHARRARD W J W and KNOWELDEN, J (1956) *Lancet* **i** 9
SUTHERLAND I N (1956) *Brit J prev soc Med* **10** 58
WEISS P and EDDS M V Jr (1946) *Amer J Physiol* **145** 587

The early electron micrographs of viruses did not give much detail of shape or contour of the particles but a great improvement was made by Williams and Wyckoff (1945) who developed the metal shadowing technique Briefly this consists of vaporizing a small quantity of metal by means of an electric current in a high vacuum so that a thin coating of metal is thrown at an angle onto the virus particle The metal coating covers the film except for the area in the shadow of the virus particle This area, uncoated by metal, is the shadow Furthermore, the shape of the shadow gives information as to the actual topography of the particle

Another important advance in the electron microscopy of viruses is the development of the ultramicrotome which enables sections to be cut measuring 200–300 Å in thickness By this means we can cut ultra thin sections of virus diseased tissues and actually see the virus inside the cell which is the next best thing to observing virus in the living cell this last is of course impossible with the electron microscope

Morphology of Plant Virus Particles

Refinements in the study of virus particles on the electron microscope, such as the technique of making carbon replicas, together with X ray studies have given considerable information on the appearance and anatomy of the tobacco mosaic virus particle (T M V) At the moment the results of the X ray work and the electron micrographs do not seem to agree precisely According to the former the particle of T M V has a helical twist whilst the latter gives the impression of a hexagonal rod Workers in California and Germany have managed to break down the particles and to observe how the protein and nucleic acid fit together The virus rod appears to be hollow and the nucleic acid without which the virus cannot multiply runs up the centre rather like the lead in a pencil Fragments of the virus particle sometimes appear as discs with a hexagonal outline and a hole in the centre which is presumably occupied by the long threads of nucleic acid

Several other plant virus particles are rod like but they differ in varying degrees from the particles of T M V Potato virus

the virus diseases of the higher animals to regard plant viruses as belonging to a separate category This is, I think, largely because some plant viruses have been obtained in a crystalline form However, there is no *a priori* reason why animal viruses should not also be crystallized provided they are small enough and can be obtained in sufficient quantities Indeed, as we shall see later, at least two animal viruses have already been crystallized

For many years virologists endeavoured to see viruses on the optical microscope and so to obtain some idea of what they looked like The best that could be done, however, was to observe agglomerates of the larger animal viruses, the so-called 'elementary bodies By means of his ultra violet light microscope Barnard was able to photograph some of the larger animal pox viruses but of course no detailed structure was visible and the resolution obtainable with ultra violet is of no use for seeing any of the plant viruses The invention and development of the electron microscope has changed all this and, with a resolution of 10 Å, the smallest virus is brought easily into range

As I have mentioned the first virus to be discovered was a plant virus that of tobacco mosaic by Iwanowski in 1892, and it was this same virus which was the first to be isolated by Stanley in 1935 and the first to be visualized by means of the electron microscope in Germany in 1939 The actual shape of the tobacco mosaic virus particle had been foretold as early as 1933 by two American workers, Takahashi and Rawlins, who examined clarified sap from a tobacco plant with mosaic under a polarizing microscope and observed the phenomenon of 'anisotropy of flow Minute rods which are contained in a flowing liquid tend to become orientated with their long axis parallel to the direction of flow rather like logs in a stream Under these conditions a liquid containing rods is doubly refractive when the direction of transmission of the incident light is perpendicular to the direction of flow This is what is meant by anisotropy of flow The observation was amply confirmed by the electron microscope and the tobacco mosaic virus was seen to be a rod measuring 350 mμ in length by 15 mμ in diameter

The early electron micrographs of viruses did not give much detail of shape or contour of the particles, but a great improvement was made by Williams and Wyckoff (1945) who developed the metal shadowing technique Briefly this consists of vaporizing a small quantity of metal by means of an electric current in a high vacuum so that a thin coating of metal is thrown at an angle onto the virus particle The metal coating covers the film except for the area in the shadow of the virus particle This area, uncoated by metal is the 'shadow Furthermore the shape of the shadow gives information as to the actual topography of the particle

Another important advance in the electron microscopy of viruses is the development of the ultramicrotome which enables sections to be cut measuring 200–300 Å in thickness By this means we can cut ultra thin sections of virus diseased tissues and actually see the virus inside the cell which is the next best thing to observing virus in the living cell this last is of course impossible with the electron microscope

MORPHOLOGY OF PLANT VIRUS PARTICLES

Refinements in the study of virus particles on the electron microscope such as the technique of making carbon replicas together with X ray studies have given considerable information on the appearance and anatomy of the tobacco mosaic virus particle (T M V) At the moment the results of the X ray work and the electron micrographs do not seem to agree precisely According to the former the particle of T M V has a helical twist whilst the latter gives the impression of a hexagonal rod Workers in California and Germany have managed to break down the particles and to observe how the protein and nucleic acid fit together The virus rod appears to be hollow and the nucleic acid without which the virus cannot multiply runs up the centre rather like the lead in a pencil Fragments of the virus particle sometimes appear as discs with a hexagonal outline and a hole in the centre which is presumably occupied by the long threads of nucleic acid

Several other plant virus particles are rod like but they differ in varying degrees from the particles of T M V Potato virus

X and the virus of cabbage black ringspot have long rather thread like particles and they lack the rigidity of T M V A recently discovered virus in *Ranunculus* sp consists of short thick rods

Many of the plant viruses have very small apparently spherical particles such as the viruses of tomato bushy stunt, turnip yellow mosaic and tobacco ringspot These viruses are some of the smallest known and measure only 25–28 milli microns in diameter (a millimicron equals a millionth of a millimetre) It is probable that they are not really spheres and Steere (1957) in California has shown that the particles of tobacco ringspot are actually polyhedral in shape

I have mentioned earlier that some plant viruses have been obtained in crystalline form and some of these virus crystals can be observed in the electron microscope One of the tobacco necrosis viruses is especially suitable for this because it crystal lizes in thin flat plates which lend themselves well to electron microscopy The micrographs show the regular packing of the virus particles on the face of the crystal

Cutting Ultra thin Sections

Provided the sections are thin enough, this technique promises to add considerably to our knowledge of the behaviour of viruses in the cells of infected organisms Material for sectioning is usually fixed in osmic acid and embedded in methacrylate since the ordinary method of wax embedding is not suitable The sections are cut on a glass knife prepared by fracturing a strip of plate glass, so that a very sharp edge is given The microtome is an instrument of high precision in which the advance of the object to the knife edge is usually accomplished by thermal expansion As the sections are cut the ribbon is floated onto a bath of alcohol and from there picked up onto the grid for transfer to the electron microscope By a refinement of this technique which cannot be described in detail here it is now possible to cut serial sections through the virus particles themselves

One of the difficulties in examining thin sections of plant tissue on the electron microscope is the interpretation of the

objects seen Obviously it is easier to distinguish a virus from the normal cell contents if it has a characteristic shape and if it is present in high concentration It is probably because T M V fulfils both these conditions that more attention has been paid to this than to other plant viruses

The normal cell cytoplasm in plants contains many small spherical particles so that it is a difficult matter to be sure of the identity of the small viruses unless some special circumstances arise which make identification possible Such circumstances would be the agglomeration of the virus particles and their arrangement in crystalline array as apparently may take place with the virus of tomato bushy stunt

Insect Viruses

Although this group had until recently, received very little attention there is accumulating now a lot of information on these very interesting viruses

We may, perhaps begin by describing briefly the different types of insect viruses as they are so far recognized The *poly hedral diseases* have been known for quite a number of years and the name is derived from the fact that certain tissues of the affected insects contain huge numbers of polyhedral or many sided crystals We shall discuss these in more detail later It has recently been shown that the polyhedral diseases consist of two quite distinct types, the *nuclear* and the *cytoplasmic* polyhedroses

The second type of insect viruses is characterized not by polyhedral crystals but by the occurrence in the tissues of numbers of extremely small granules from which the name *granular* diseases or granuloses is derived

In the third type of insect virus no crystals granules or similar intracellular inclusions are involved and the virus par ticles occur freely in certain tissues

Most of these viruses attack only the larval forms of the in sects As a whole the insect viruses are very suitable for electron microscopy since they are of rather large size occur in high concentration in the tissues and are fairly easily observed inside affected cells

Polyhedral Viruses

In considering the nature of the polyhedral crystals, it is important to understand the fundamental difference between these and a virus crystal A polyhedral crystal is a protein crystal in which virus particles are enclosed, a virus crystal is a nucleoprotein consisting only of virus This difference is easily seen by comparing Plate XXX, Figure 3, which is an electron micrograph of a section through a nuclear polyhedral crystal showing the rod shaped virus particles enclosed within, with Plate XXXII, Figure 5, which is also an electron micrograph of a section through an insect virus crystal composed only of near spherical particles

We have said that there are two types of polyhedral viruses and they can easily be differentiated by the following simple procedure A smear is made of the body contents of the affected insect larva and stained with Giemsa solution The nuclear polyhedral crystals remain unstained and show up against the stained background, the *cytoplasmic polyhedra, on the other hand, take up the stain* readily

There are two methods by which the presence of the virus inside the polyhedral crystal can be determined, one is by treating the crystals with dilute sodium carbonate and the other by cutting thin sections of the crystals It is rather interesting to find that the nuclear and the cytoplasmic polyhedra differ from each other in some important characteristics When treated with weak alkali and viewed on the electron microscope, the nuclear polyhedra will be seen to have dissolved, leaving behind a membrane which enclosed the crystal and in which the rod shaped virus particles can be plainly seen The cytoplasmic polyhedra on the other hand, when similarly treated, react differently There is no enclosing membrane and the crystals dissolve only partially, leaving behind a matrix filled with round holes In some cases the spherical virus particles which occupy these holes can be seen but more frequently the virus itself is dissolved For this type of polyhedra and for a special type of nuclear polyhedra from a fly larva, *Tipula paludosa,* it is more satisfactory to cut thin sections of the crystal itself This is

Polyhedral Viruses

In considering the nature of the polyhedral crystals, it is important to understand the fundamental difference between these and a virus crystal A polyhedral crystal is a protein crystal in which virus particles are enclosed, a virus crystal is a nucleoprotein consisting only of virus This difference is easily seen by comparing Plate XXX, Figure 3, which is an electron micrograph of a section through a nuclear polyhedral crystal showing the rod shaped virus particles enclosed within, with Plate XXXII, Figure 5 which is also an electron micrograph of a section through an insect virus crystal composed only of near-spherical particles

We have said that there are two types of polyhedral viruses and they can easily be differentiated by the following simple procedure A smear is made of the body contents of the affected insect larva and stained with Giemsa solution The nuclear polyhedral crystals remain unstained and show up against the stained background, the cytoplasmic polyhedra, on the other hand, take up the stain readily

There are two methods by which the presence of the virus inside the polyhedral crystal can be determined, one is by treating the crystals with dilute sodium carbonate and the other by cutting thin sections of the crystals It is rather interesting to find that the nuclear and the cytoplasmic polyhedra differ from each other in some important characteristics When treated with weak alkali and viewed on the electron microscope, the nuclear polyhedra will be seen to have dissolved, leaving behind a membrane which enclosed the crystal and in which the rod shaped virus particles can be plainly seen The cytoplasmic polyhedra, on the other hand, when similarly treated, react differently There is no enclosing membrane and the crystals dissolve only partially leaving behind a matrix filled with round holes In some cases the spherical virus particles which occupy these holes can be seen but more frequently the virus itself is dissolved For this type of polyhedra and for a special type of nuclear polyhedra from a fly larva *Tipula paludosa*, it is more satisfactory to cut thin sections of the crystal itself This is

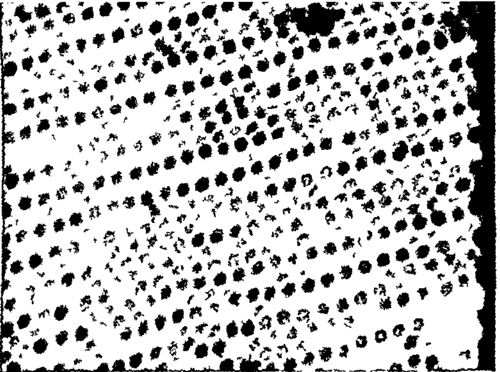


FIG 5 Ultra thin section through a crystal similar to that shown in Fig 4 note the regular packing of the virus particles compare Fig 3 (× 28 000)
(After Williams and Smith 1957)



FIG 3 Section through a nuclear polyhedral crystal from the larva of a fly *Tipula paludosa* note the virus rods embedded in the crystal matrix compare Fig 5 (× 29 000)

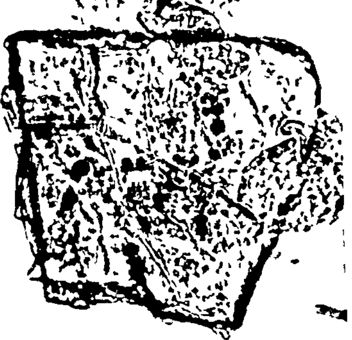


FIG 4 A virus crystal from a larva of *T paludosa* infected with a cytoplasmic virus without intracellular inclusions photographed on the optical microscope (× 19 0)
(After Williams and Smith 1957)

difficult, however, because of the hardness of the crystal Plate XXXI, Figure 3 is a section through a nuclear polyhedral crystal from the larva of *T paludosa* showing the virus rods contained within

The following, then, are the characteristics of the two kinds of polyhedral disease, in the nuclear diseases the crystal dissolves completely in weak sodium carbonate leaving behind a membrane containing rod shaped virus particles The virus develops in the nuclei of cells in the skin, tracheae and blood and sometimes apparently in the mitochondria (Plate XXXIII Figure 6) In the cytoplasmic diseases there is no membrane round the crystal, which dissolves only partially leaving behind a framework containing circular holes The virus particles are spherical or near spherical and the virus always develops in the cytoplasm of cells of the gut In addition the cytoplasmic polyhedra stain with Giemsa solution thus differing from the nuclear polyhedra which do not stain

In the second group of insect virus diseases the *granuloses* we can apply the same sodium carbonate technique to dissolve the granules This reveals the curious tendency of insect viruses to be enclosed within capsules and membranes The granule partially dissolves and flattens exposing a rod shaped body within which however is not the virus particle but another capsule The particle is contained within the rod shaped capsule and is also rod shaped further treatment with sodium carbonate releases the virus rod from the capsule and will also dissolve the substance of the virus rod itself leaving behind as a ghost the intimate membrane which contains the virus material

We come now to the last group of insect viruses in which there are no polyhedra, granules and other intracellular inclusions Very few of this kind of virus have been described and attention here will be paid only to an interesting virus recently isolated from a fly larva *Tipula paludosa* This is the same species of larva from which a polyhedral virus has been obtained

This new virus is rather large and it appears spherical but is actually polyhedral in shape it multiplies in the cytoplasm of the fat body cells It occurs in very high concentration in the insect and as much as 25 per cent of the actual body weight of



FIG 6 Section through three apparent mitochondria in a larva of *T paludosa* infected with a nuclear polyhedral virus note the virus particles surrounded by a membrane in two of the three apparent mitochondria the centre one appears free of virus ($\times$ 45 000)

The other virus is not regarded by some virologists as belonging to the virus category but is put in a separate class, the Rickettsiae These agents are certainly considerably larger than the average virus and may perhaps be regarded as a link between the viruses proper and the bacteria Sections of certain Rickettsiae show a limiting membrane and an intermediate zone of granular material, there is also a space containing a central body A somewhat similar state of affairs is seen in sections of vaccinia and herpes viruses where there is also a central body and a double membrane

REFERENCES

IWANOWSKI D (1892) *Bull Acad imp Sci St Petersb* N S 5 67

STANLEY W M (1935) *Science* N S 81 644

STEERE R L (1957) *J biophys biochem Cytol* 3 45

TAKAHASHI W N and RAWLINS T E (1933) *Science* N S 77 26

WILLIAMS R C and WYCKOFF R W G (1945) *Proc Soc exp Biol Med* 58 267

WILLIAMS R C and SMITH K M (1957) *Nature Lond* 179 119–120

the larva is converted into virus When a drop of the body fluid from an infected larva is observed on the electron microscope it appears as a sheet of five and six sided particles arranged in a regular manner Because of its extremely uniform size and high concentration in the insect, this virus has proved very suitable for studies on purification and crystallization Professor R C Williams and the writer (1957) have now succeeded in crystallizing this virus and a crystal photographed on the optical microscope is shown in Plate XXXI, Figure 4 When such a crystal is sectioned and observed on the electron microscope the close packing of the particles in the crystal is such as can be seen in Plate XXXII Figure 5 It will be noticed in this electron micrograph that the particles do not appear to be in contact This is thought to be due to the fact that each virus particle is enclosed in a membrane which is very much less dense than the central portion of the virus and so is not visible, though each membrane is probably in contact with its neighbour

This same virus also exhibits some extremely interesting optical properties When the virus is sedimented on the centrifuge at 10,000 r p m , the pellet exhibits fascinating colour changes By transmitted light the pellet appears an orange or amber colour By reflected light it has an iridescent, turquoise appearance Within the pellet may be seen small regions reflecting the incident light quite brilliantly, giving the entire pellet the appearance of an opal The particles are separated from each other in the crystal and this gives the light scattering which produces the turquoise effect

Some Viruses Affecting Higher Animals

For the sake of comparison with the viruses affecting plants and insects, it may be worth while to mention briefly two viruses from man One of these that of poliomyelitis has been studied at Berkeley California and was one of the first, if not the first animal viruses to be obtained in crystalline form It is a very small virus and seen on the electron microscope has a close superficial resemblance to a small plant virus like that of tomato bushy stunt

L C THOMSON
*The Physiological Basis of Visual Sensation

J M MACKINTOSH
The Contribution of Science to the Practice of Health in the First Quarter of the Twentieth Century

ALICE M STEWART
Methods of Research in Social Medicine

G P CROWDEN
Environmental Factors in Work

DONALD HUNTER
Methods of Research in Industrial Medicine

H D KAY
Recent Light on Mammary Function

E C DODDS
Research on Ageing

A HADDOW
Carcinogenesis

W V MAYNEORD
Physical Techniques in the Medical Applications of Ionizing Radiations

J F LOUTIT
Biological Effects of Radiation

O G EDHOLM
*The Effects of Haemorrhage on the Cardiovascular System in Man

M L ROSENHEIM
*Lability of Blood Pressure

J H M MICHAEL
Cardiac Output in Man (*film*)

1952–53

J Z YOUNG
The Influence of Language on Medicine

SIR FREDERICK BARTLETT
The Nature and Place of Thinking in Medicine

SIR JAMES SPENCE
The Methodology of Clinical Science

BRADFORD HILL
The Statistical Approach

C G DOUGLAS
Control of Respiration

SIR B H C MATTHEWS
Life at High Altitudes

W K STEWART
*The Physiological Effects of Gravity

F J W ROUGHTON
The Kinetics of Rapid Chemical and Biological Reactions

D KEILIN
Metal Catalysis and Intracellular Respiration

J T RANDALL
Biophysical Studies of Connective Tissue

A H T ROBB SMITH
*The Functional Significance of Connective Tissue

E J KING
Silicosis

MATTHEW STEWART
Pulmonary Asbestosis

HONOR B FELL
Organ Culture in Biological and Medical Research

P C C GARNHAM
*The Life History of the Malaria Parasite

J C WHITE
*Human Haemoglobins

SIR ALAN DRURY
Uses and Applications of Human Blood Plasma Fractions

A S PARKES
*Preservation of Living Cells at Low Temperatures

P L MOLLISON
*The Life span of Red Blood cells

A NEUBERGER
*Biochemical Genetics

C E DENT
*Chromatography in the Study of Amino-Acid Metabolism

ADRIAN ALBERT
*Selective Toxicity (*2 lectures*)

COMPLETE LIST OF LECTURES

* *Lectures included in the published volumes*

1951–52

E ADRIAN
*The Scientific Approach to Medical Research

H DINGLE
The Philosophy of Science

W T ASTBURY
Studies by X Ray Analysis Electron Microscopy and Supporting Techniques of the Structure of Biological Macromolecules and the Tissues formed from them

G R CAMERON
**Tissue Responses to Injury*

D VAN SLYKE
*Studies of Normal and Pathological Physiology of the Kidney

R A PETERS
Biochemical Function of Vitamin B_1

S J COWELL
Nutritional Science and the Feeding of Populations

J YUDKIN
*Nutritional Assessment of the Individual

D D WOODS
Folic Acid and Vitamin B_{12} in the Metabolism of Micro-organisms

B C J O KNIGHT
Aspects of Bacteriostasis

A A MILES
*Some Aspects of Antibacterial Immunity

A W DOWNIE
*Antibodies and Immunity to Virus Infection

F C BAWDEN
Current Knowledge on Nature of Viruses

S P BEDSON
**Viruses as the Causes of Diseases*

F G YOUNG
*Adrenal Hormones and ACTH

SIR ROBERT ROBINSON
Chemical Aspects of Antibiotics

J M TANNER
*Growth of the Human at the Time of Adolescence

W D NEWCOMB
Bone Growth and Repair

L J WITTS
The Rate and Materials of Blood Formation

L J WITTS
Alimentary Factors in Blood Formation

R G MACFARLANE
*Blood Coagulation in Theory and Practice

SIR CYRIL BURT
The Psychology of Personality

SIR GEOFFREY JEFFERSON
*On the Organization of Cortical Mechanisms

G L BROWN
Involuntary Nervous System *(2 lectures)*

LORD STAMP
*The Action of Bacterial Enzymes on Immunizing Antigens

H B MAITLAND
*Antiviral Immunity

C C CHESTERMAN
*Antimalarial Drugs

P M D ARCY HART
Chemotherapy of Tuberculosis

R R RACE
Blood Groups

J V DACIE
Congenital Haemolytic Anaemia

J H KELLGREN
*The Supporting System and its Disorders

G M BULL
*Regulation of Body Water

1954-55

A V HILL
*Why Biophysics?

B S PLATT
*Protein Malnutrition

J D BOYD
The Importance of Detail in the Clinical Anatomy of the Autonomic Nervous System

D WHITTERIDGE
*The Effects of Visceral Distension

P M DANIEL
Some Features of the Peripheral Circulation and Vascular Bed

W T J MORGAN
*The Chem cal Basis of Blood Group Specificity in Man

G M WILSON
*The Electrolyte and Metabol c Response to Trauma

J D JUDAH
*Enzymes n Injury

J P BULL
*Shock from Burns

A G EVERSON PEARSE
*Histochemistry and its Application to the Bas c Sciences

SIR VICTOR E NEGUS
*Comparative Anatomy of the Larynx

R I S BAYLISS
Factors Influencing Adrenocortical Activity in Health and Disease

L MARY PICKFORD
*Release and Action of Posterior Pituitary Hormones

ROSALIND PITT RIVERS
Thyroid Hormones

F BERGEL
*Some Chemical Aspects of Abnormal Growth

A HADDOW
Theory and Application of the Nitrogen Mustards

SIR RUDOLPH PETERS
Medical Significance of Biochemical Lesions

DOUGLAS McCLEAN
*Substances that Increase Tissue Permeability and their Relation to Infection and Fertilization

D G MELROSE
*Cooling of the Whole Organism

JANET VAUGHAN
*Radiation Effects on Bone

E F GALE
Actions of Antibiotics on Bacteria

P B MEDAWAR
Tissue Transplantation Immunity

J R SQUIRE
*Correlation between Laboratory and Cl nical Findings in Hypersensitivity

C H GRAY
*The Chemistry of the Porphyrias

T S WORK
Protein Biosynthesis

J N DAVIDSON
*Nucleoproteins in Cell Structure

J H GADDUM
*The Effects of Alcohol

A C FRAZER
*Fat Metabolism

SIR ALEXANDER FLEMING
*Recent Progress in Antibiotics

WILSON SMITH
*Virus Adaptability in Relation to Human Disease

P R PEACOCK
*Carcinogenesis

G F MARRIAN
*The Metabolism of the Adrenocortical Hormones

S J FOLLEY
The Pituitary Gland and Reproduction

A ST G HUGGETT
*The Physiology of Parturition

W S FELDBERG
*The Physiology of the Autonomic Nervous System

W D M PATON
*The Principles of Ganglionic Block

R H S THOMPSON
*Cholinesterases and Anti Cholinesterases

H J SEDDON
Certain Aspects of Nerve Repair

J D BOYD
Development of the Heart in Relation to Congenital Heart Disease

A HEMINGWAY
Dynamics of the Heart Beat

G W PICKERING
The Natural History of Essential Hypertension

E P SHARPEY SCHAFER
Causes of Hypotension in Man

1953-54

SIR HENRY DALE
Scientific Method in Medical Research

R W RUSSELL
*Experimental Psychopathology

G POPJÁK
*Biological Synthesis

J H BURN
*Acetylcholine and the Maintenance of the Cardiac Rhythm

F G YOUNG
*The Growth Hormone of the Anterior Pituitary Gland

E C AMOROSO
The Biology of the Foetal Membranes and the Placenta

P M F BISHOP
*The Physiological Actions of the Sex Hormones

E J KING
*Acid and Alkaline Phosphatase in Disease

R L M SYNGE
Principles of Chromatography

W T ASTBURY
Some Recent Ideas about the Proteins and other Biological Macromolecules

G W HARRIS
*Stress and Thyroid Activity

H L SHEEHAN
The Vascular Lesions of Pituitary Necrosis

H E SIGERIST
*Science and History

W D M PATON
*Anticholinesterases

R M B MACKENNA
*The Scientific Approach to Dermatology

E A CARMICHAEL
*Hemispherectomy and the Localization of Function

G R CAMERON
*Tissue Repair

D D REID
The Design of Clinical Experiments

J B S HALDANE
*The Genetics of Some Biochemical Abnormalities

E BOYLAND
*The Chemotherapy of Cancer

L P GARROD
*Causes of Failure in Antibiotic Therapy

A A MILES
*Reactions to Bacterial Invasion

F H C CRICK
The Structure and Function of Small Viruses

V P WHITTAKER
*The Metabolism of Choline Esters

MARY BARBER
*Resistance of Staphylococci to Antibiotics

E C AMOROSO
Endocrinology of Pregnancy

W J HAMILTON
Multiple Pregnancy

W J DEMPSTER
*Homotransplantation of Organs

C G ROB
*Arterial Substitutes

SIR RUSSELL BROCK
The Relation between the Embryology and Morphology of the Heart and Cardiac Surgery

P WOOD
The Physiological Basis of Clinical Signs in Heart Disease

J C GILSON
*Tests of Lung Functions an Assessment of their Usefulness

D G MELROSE
*Principles of Heart Lung Machines

C LONG
*Phospholipids

SIR SOLLY ZUCKERMAN
Biological Systems

G HADFIELD
*The Role of the Endocrine System in Breast Cancer

A J LEWIS
*Social Psychiatry

E G T LIDDELL
*Cajal and Sherrington

H McILWAIN
*Neurochemistry

E J ZAIMIS
*Factors Influencing the Action of Neuromuscular Blocking Substances

C A KEELE
*Causes of Pain

N B MYANT
*Biliary Excretion of Thyroid Hormone

H M SINCLAIR
*Vitamins in Nutrition

H J SEDDON
*Recent Work on Paralytic Poliomyelitis

W J WHELAN
The Synthesis and Degradation of Polysaccharides

A L GREENBAUM
*The Control of Fat Metabolism

HONOR B FELL
*The Physiology of Skeletal Tissue in Culture

J B DUGUID
Arterial Occlusion

H GRÜNEBERG
The Localization of Inherited Disease in the Body

R A KEKWICK
*The Plasma Proteins

C H ANDREWES
*The New Look in Virus Research

1955-56

SIR WILFRID LE GROS CLARK
*Hypothesis and Speculation in Scientific Research

M M SWANN
Physiology of Mitosis

H G DAVIES
*The Use of the Interference Microscope in Biological Research

O G EDHOLM
*The Effects of Cold on Man

A C DORNHORST
*Physiology of the Lower Oesophagus and Cardia

R E TUNBRIDGE
*Observations on the Structure of Connective Tissue Fibres

N H MARTIN
*Some Aspects of Protein Diseases

R D HARKNESS
*Metabolism of Collagen

J C MCCLURE BROWNE
*Isotopes in the Study of Problems of Pregnancy

G S DAWES
*Physiological Effects of Anoxia in the Foetal and Newborn Lamb

J WALKER
*The Oxygen Environment of the Foetus

D A SLOME
*Physiology of Nasal Circulation

AUDREY U SMITH
*Experimental Hypothermia in Animals

J N HUNT
*The Investigation of Gastric Digestive Function in Man

R E DAVIES
Biochemical Aspects of Gastric Secretion

H A KREBS
*Steering of Metabolic Processes

W F J CUTHBERTSON
*The Nutrition of Micro-Organisms

R A MORTON
*Vitamin A

K M RUDALL
*Protein Ribbons and Sheets

M D MILNE
*Renal Control of Acid Base Balance

D R WILKIE
*Living Muscle

S V PERRY
*Proteins in Muscular Contraction

J F STOKES
*The Treatment of Hepatic Coma

LORD ROTHSCHILD
Fertilisation

T R R MANN
Mammalian Semen Composition and Function

E S HORNING
*Some Anomalies in Endocrine Carcinogenesis

K C RICHARDSON
Problem of Mammary Growth and Structure

M W GOLDBLATT
*Industrial Toxicology

J M BARNES
*The Elucidation of Toxicity

J A V BATES
*Observations on the Cortical Motor Areas in Man

J F LOUTIT
*Recovery from Lethal Effects of Ionising Radiation

1956-57

SIR JOHN COCKCROFT
*Biological Significance of Atomic Energy

K M SMITH
*The Electron Microscope in the Study of Viruses